AGARD-CP-541

AGARD-CP-541

AGARD

ADVISORY GROUP FOR AEROSPACE RESEARCH & DEVELOPMENT
7 RUE ANCELLE 92200 NEUILLY SUR SEINE FRANCE

AGARD CONFERENCE PROCEEDINGS 541

Virtual Interfaces: Research and Applications

(Les Interfaces Virtuelles
entre Recherche et Applications)

Papers presented at the Aerospace Medical Panel Symposium held in Lisbon, Portugal, 18th–22nd October 1993.

NORTH ATLANTIC TREATY ORGANIZATION

Published May 1994

Distribution and Availability on Back Cover

The Mission of AGARD

According to its Charter, the mission of AGARD is to bring together the leading personalities of the NATO nations in the fields of science and technology relating to aerospace for the following purposes:

- Recommending effective ways for the member nations to use their research and development capabilities for the common benefit of the NATO community;
- Providing scientific and technical advice and assistance to the Military Committee in the field of aerospace research and development (with particular regard to its military application);
- Continuously stimulating advances in the aerospace sciences relevant to strengthening the common defence posture;
- Improving the co-operation among member nations in aerospace research and development;
- Exchange of scientific and technical information;
- Providing assistance to member nations for the purpose of increasing their scientific and technical potential;
- Rendering scientific and technical assistance, as requested, to other NATO bodies and to member nations in connection with research and development problems in the aerospace field.

The highest authority within AGARD is the National Delegates Board consisting of officially appointed senior representatives from each member nation. The mission of AGARD is carried out through the Panels which are composed of experts appointed by the National Delegates, the Consultant and Exchange Programme and the Aerospace Applications Studies Programme. The results of AGARD work are reported to the member nations and the NATO Authorities through the AGARD series of publications of which this is one.

Participation in AGARD activities is by invitation only and is normally limited to citizens of the NATO nations.

The content of this publication has been reproduced
directly from material supplied by AGARD or the authors.

Published May 1994

ISBN 92-835-0746-0

Printed by Specialised Printing Services Limited
40 Chigwell Lane, Loughton, Essex IG10 3TZ

Preface

The overall effectiveness of aerospace systems can be greatly improved by more efficient use of human performance and human decision making. Most aerospace systems that involve a human and a responsive machine appear limited by the design of the interface between them. These interfaces support the human's situational awareness and provide interpreted command and control for mechanistic implementation.

Recent advances in technologies for information display and sensing of human movements, combined with computer based models of natural and artificial environments, have led to the introduction of so-called virtual interfaces. Virtual interfaces offer increased flexibility and naturalness, so are considered for use in several domains including aviation, training, design, simulation and robotics.

Papers presented at this symposium considered issues of research and application in virtual interfaces broadly defined. Issues of technology integration for system development were considered separately from issues of movement monitoring or sensory display. Issues of human performance measurement were presented in the context of both research and application. A description of systems in engineering development for cockpit and for telesurgery was also presented.

Préface

L'efficacité globale des systèmes aérospatiaux peut être considérablement améliorée par l'exploitation plus judicieuse des performances humaines et l'emploi effectif de la prise de décision humaine. La plupart des systèmes aérospatiaux qui mettent en présence un être humain et une machine interactive semblent être limités par le type d'interface qui les réunit. Ces interfaces renforcent la perception de la situation par l'opérateur humain et fournissent des éléments interprétés de commandement et de contrôle pour application mécanique.

Les progrès réalisés récemment dans le domaine des technologies de l'affichage des données et de la détection des mouvements humains, alliés aux modèles informatisés des milieux naturels et artificiels, ont conduit à la mise en place d'interfaces dites virtuelles. Les interfaces virtuelles offrent plus de souplesse et de naturel et elles sont donc envisagées pour les domaines tels que l'aviation, la formation, la conception et la robotique.

Les communications présentées lors de ce symposium examinaient certains sujets de recherche et de leurs applications dans le domaine des interfaces virtuelles dans le sens large du terme. Les questions concernant l'intégration des technologies aux fins du développement des systèmes ont été considérées séparément des questions de suivi des mouvements ou de l'affichage sensoriel. Les questions concernant l'évaluation des performances humaines ont été présentées dans le double contexte de la recherche et des applications. Une description des systèmes destinés à l'habitacle et à la téléchirurgie et actuellement au stade de développement de l'ingénierie a également été fournie.

Aerospace Medical Panel

Chairman: Prof. Dr Med. L. Vogt
Institut für Flugmedizin
DLR
Postfach 90 60 58
Linder Höhe
D-51147 Köln 90
Germany

Deputy Chairman: Colonel P. Vandenbosch
VSM
Director Aeromedical Services BAF
Kwartier Koning Albert 1er
Raketstraat 70
B-1130 Brussels
Belgium

TECHNICAL PROGRAMME COMMITTEE

Dr J.F. Tangney (**Chairman**)
AFOSR/NL
110 Duncan Ave.- Suite B1 15
Bolling AFB
DC 20332-0001
United States

Dr S. Hart
Chief, Rotorcraft Human Factors
NASA Ames Research Center
Moffett Field
CA 94035
United States

Dr K.R. Boff
AL/CFH
Bldg 248 - 2255 H. Street
Wright-Patterson AFB
OH 45433-7022
United States

Dr A. Léger
Sextant Avionique
Rue Toussaint Castros
BP 91
33166 Saint Médard en Jalles
France

Ms J. Davies
British Aerospace Defence Ltd
Military Aircraft Division
Farnborough Aerospace Centre
PO Box 87
Farnborough
Hants GU14 6YU
United Kingdom

Dr J. Smit
Head, Research and Development
National Aerospace Medical Center
Kampweg 3 - PO Box 22
3769 ZG Soesterberg
The Netherlands

Local Coordinator
Major N.P. Ribeiro
Base do Lumiar
Centro de Medica Aeronautica
1600 Lisbon
Portugal

Host Nation Coordinator
Col. A. Entrudo
Director - Direccao de Saude
Forca Aerea Portugesa
Av. Leite de Vasconcelos
Alfragide - Amadora
Portugal

PANEL EXECUTIVE

Major R. Poisson, CAF

Mail from Europe:
AGARD–OTAN
Attn: AMP Executive
7, rue Ancelle
92200 Neuilly-sur-Seine
France

Mail from USA and Canada:
AGARD–NATO
Attn: AMP Executive
PSC 116
APO AE 09777

Tel: 33(1) 47 38 57 60
Telex: 610176 (France)
Telefax: 33 (1) 47 38 57 99

Contents

TECHNICAL EVALUATION REPORT

by

John F. Tangney, Ph.D.
Air Force Office of Scientific Research
Washington DC USA 20332-0001

1. INTRODUCTION

The Aerospace Medical Panel held a Symposium on "Virtual Interfaces: Research and Applications" at facilities of the Portugese Air Force located in Lisbon, Portugal, 18-22 October 1993. Twenty papers were presented along with an invited address and three video-taped demonstrations of interface technologies, and some round table discussion of major issues for research and development. Papers represented contributions by five NATO countries, with ninety registrants in attendance representing twelve NATO countries.

2. THEME

At the 72nd Business Meeting of the Aerospace Medical Panel, held October 1991 in Rome Italy, approval was obtained for a symposium to present the current state of research and development in synthetic interfaces with the goal of informing system designers who might be considering the use of such interfaces in aerospace environments.

Discussion of this topic at previous Business Meetings revealed a broad interest in the topic of virtual interfaces, but too few sustained efforts in research or development across the NATO countries to support lengthy consideration of the lessons learned or the research findings that might be used to inform efforts of other member countries.

By fall of 1991, however, efforts using virtual interfaces were underway in several NATO countries. It also became increasingly clear that the multi-disciplinary nature of the research and the number of options possible for implementing any interface design were sufficiently large that a symposium for reports of progress would benefit all. A greater degree of coupling between these efforts then became a secondary goal of the planned symposium to consider virtual interfaces in aerospace application domains.

3. PURPOSE AND SCOPE

The interface between humans and machines is changing dramatically in aerospace occupations with the introduction of new sensing technologies that permit continuous monitoring of human movements, and new display technologies that can provide substitutes for the normal experiences of vision, hearing, touch and other senses. When used in combination, these technologies can be used to create a "virtual interface" for human operators through the closed loop computerized control of their sensory experience.

Effective implementation of virtual interfaces presents a number of challenges to basic and applied research scientists. System designers must select components from a varied assortment of hardware and software for each job implemented. The specifications for these components can vary widely. Human factors scientists must specify the costs and benefits, in terms of human performance, of using these technologies for specific work environments and must adapt these technologies to different tasks. Basic researchers are challenged to develop more complex models of human performance that can be used to constrain the design process.

Papers were solicited on three broad topics of virtual interface: (1) the sensing of human movement and posture, (2) the display of information to human operators, and (3) the issues of system integration. Submitted papers were reviewed by the technical program committee, as approved by the Aerospace Medical Panel, consisting of Dr. K. Boff (US), Dr. J. Davies (UK), Dr. S. Hart (US), Dr. A.

Leger (FR), and Dr J. Smit (NE). This committee was assisted by a NATO-provided advisor, Dr N. Durlach (US).

4. SYMPOSIUM PROGRAM

The symposium included a keynote address on the research agenda for virtual interfaces delivered by Professor Kalawsky of the UK, and four technical sessions: (1) System Integration I, chaired by Dr. J. Davies (UK); (2) System Integration II, chaired by Dr. J. Smit (NE); (3) Sensory Technology plus Evaluation, chaired by Dr. A. Leger (FR) and LCdr D. Dolgin (US); and (4) Human Performance Issues, chaired by LCol S. Porcu (IT) and Dr. K. Boff (US).

Video-taped demonstrations were also shown, on the use of virtual interfaces for telesurgery, architectural design, and telerobotics. A general discussion capped the meeting in an attempt to reach consensus on conclusions and recommendations.

5. TECHNICAL EVALUATION

In his keynote address, Professor Kalawsky surveyed the domain of virtual interfaces, including issues of definition, application, research, and business decisions affecting progress in the field. He proposed adopting a three part definition that includes computer-based models of an environment (autonomy) combined with an ability for human interaction (interaction) done in a way that supports natural modes of human interaction (presence). In this survey of application domains and research issues, he noted the relative maturity of visual display devices (for example) compared to other interface technologies of tactile displays or haptic sensing (for example). He concluded by emphasizing the need for truly collaborative multidisciplinary work at a system level to afford progress in the field.

5.1 System Integration 1

This session included descriptions of systems for training, medicine and general graphical user interfaces.

A broad overview of interface design was provided by Nilan (paper #1), emphasizing an efficiency criterion to distinguish the value of one interface design from another. Without mentioning specific ways that efficiency of two or more designs might be measured and compared, Nilan pointed to examples where guidlines for design were extracted from social psychological research and cognitive research to greatly improve the speed and accuracy of performance using the redesigned interfaces. Such general guidlines can be applied in different task domains by extensive first use of user surveys, for example. The paper describes one example involving the redesign of a graphical user interface that more naturally matches information requirements to operator inputs needed to gather the information.

The Kennedy (paper #2) report on motion sickness is reported later in the section on Human Performance.

Medical applications for virtual environments were surveyed by Dumay and Jense (paper #3), who presented a taxonomy of application domains (including medical education, training, surgery and radiology) that might take advantage of virtual interfaces. While generally optimistic about the potential for virtual systems, Dumay mentioned a current lack of commercial availability and suggested that this may be due, in part, to a lack of high precision devices for display (visual, tactile, and force feedback) and suitable models of medical objects.

A system for training air traffic controllers was described in a paper by Marque and colleagues (paper #4). The system is not fully virtual, but relies extensively on voice recognition and artificial speech, combined with some expert systems, to replace the human teacher in a simulation environment. Performance of the speech recognition system is described in some detail, and provides an example of the ways that multi-sensory processing might be used to add value to existing training regimes.

5.2 System Integration 2

The combined use of multi-sensory virtual interface for pilots was described by Barbier and colleagues (paper #5), in an effort to increase the naturalness of the human machine interface by including vision, speech and gestural devices in a single system. Several experiments were described using the system in which the speed

of decision-making or response execution was measured for each of several interface design options.

Novel approaches to system design were outlined by Wierda and colleagues (paper #6) and described in the context of a virtual trainer for vehicular control (driving). An approach based on cognitive engineering can intuitively be shown to provide a means for separating the design goals of a system from the implementation strategies for meeting those goals. For example, a complete model of the decision making during vehicular control could be used to automate certain portions through systems of expert aiding or through improvements to the human interface.

An apparatus for measuring human performance under acceleration was described by Chelette and colleagues (paper #7) and used to assess the magnitude of the G-excess illusion under conditions of several constant G(z) loads and static head yaws. The apparatus, implemented on a centrifuge, includes helmet mounted virtual visual displays, combined head tracking, and a device for recording hand position to indicate perceived spatial attitude. The device permits experimental manipulation of the coupling between visual and vestibular inputs to human operators. Primary results concern the sensitivity of illusory tilt (pitch and roll) to head position under G-load.

A prototype and testbed for a virtual cockpit was described by Ineson (paper #8), with emphasis on the implementing hardware and the display formats for primary flight control. The system features selectible display options (e.g. visual stereo, variable terrain features) and establishes an apparatus for assessment of primary design options and human factors issues in virtual flight control. One interesting human factor issue concerns the possible confusion of head and aircraft attitude change (given some transport delay in visual image generation).

A product for high-realism in virtual visual displays was presented by Grimsdale (paper #9) with integrated hardware and software that support near real-time updates of realistically rendered complex objects (such as an automobile imaged with glints, reflections, shadows and texturing).

5.3 Sensory Technology plus Evaluation

Performance with a novel virtual hand controller was compared with a standard joy-stick controller in a preliminary report of experiments by Eggleston (paper #10). Primary findings concern rough equivalence between controllers in a task of single axis continuous tracking. Methodological issues were also raised concerning techniques of comparing devices whose parametric descriptions may not be valid (in this case, in terms of underlying kinematics).

A device for measuring point of gaze was used in an experiment reported by Zon and colleagues (paper #11) to assess the increased accuracy in reporting visual detail as dwell time of fixations increase. The device incorporates an infrared camera for tracking eye position (with respect to head) using the bright pupil method, and a six degree of freedom head tracking module. Calibration of the device and its specifications are described together with the data mentioned earlier.

A second device for measuring point of gaze was reported by Stampe (paper #12) and used to demonstrate how calibrative functions could be performed continuously and adaptively in environments where the positions of some targets for eye-movements are known. Human factors issues of keyboard type visual displays were also discussed in the context of matched resolution both spatial (between precision of eye position measures and spacing of targets) and temporal (the optimal fixation dwell time to indicate selection of the target rather than search). Experiments demonstrate that, as an input device, visual selection could provide bandwidth sufficient for several special purposes.

A device for head and eye position monitoring, installed in a centrifuge, was described by Sandor and colleagues (paper # 13). Eye movements were recorded using a corneal reflex technique, with head movements recorded by tracking helmet mounted infrared emitters. This technique does not use magnetic field sensing and so is well suited to the environment of a centrifuge. Preliminary results are reported in the tracking of continuous and saltatory visual

targets with combined head and eye movements under (up to) 6 G(z).

Extensive data were reported on the relative performance of two commercially available position sensing devices by Williams (paper #14). Measured factors include stability, noise, cross-talk, linearity, and distortion caused by metallic interference. Some detailed discussion of the devices was presented together with a description of the evaluation scheme. Differences are reported that would affect the selection of one device over another for different operating environments or different design specifications.

5.4 Human Performance Issues

A newly constructed centrifuge based flight simulator was described by Lawson and colleagues (paper #15). In a series of experiments to examine ameliorating effects of visual cues in disorienting situations of head movements under G fields, in this case produced by rotating supine subjects about a vertical axis passing through the head. Primary results concern findings similar to those found with other axes of rotation: in the dark, weaker disorientation during accelleration than during constant velocity; with visual stimuli, these effects are attenuated. Results are discussed in terms of potential problems with centrifuge based flight simulations.

Two papers presented data suggesting that stomach awareness may develop when using virtual environments. Kennedy and colleagues (paper #2) described a scoring system used to assess the degree and the nature of motion sickness. Data support the notion that three types of effect (nausea, disorientation, and oculo-motor) may be produced. The pattern of effects appear stable at different installations, suggesting that each effect may be produced by a specific failure to provide fidelity in implementation.

The second paper concerning motion sickness like reports was presented by Regan (paper #16), who described the frequency and magnitude of malaise in approximately 150 subjects. After twenty minutes, roughly five percent withdraw from the experiment due to malaise, with roughly half showing malaise at level two or greater. Partially successful coping strategies are also reported. These include reductions in the frequency, speed and magnitude of head movements.

Retinal image quality may be degraded under some conditions of virtual imagery. Kotulak and colleagues (paper #17) report non-optical factors can determine monocular visual accommodation in a fraction of viewers if the optical distance and physical distance of seen objects differs by a large factor (as can happen, for example, with vision aiding devices). Most virtual systems are binocular, however, so these results may not extend to the majority of those systems.

A software suite of tools for generating virtual imagery on personal computers was described and offered by Stampe and Grodski (paper #18). The software provides the capability for wire drawings, and stereo displays at reasonable frame rates. The software includes a set of mathematical routines that provide multiple viewpoints of the same virtual objects for multiple viewers.

Issues in three dimensional audio for virtual interfaces was described by Pellieux and colleagues (paper #20), using a facility for measuring individual Head-Related Transfer Functions (HRTF's) that describe the spectral weighting of sound sources at different locations. Experiments are reported on accuracy of locating virtual auditory sources in three space. Individual differences are reported. Greater accuracy in azimuth versus elevation is reported. The use of audio cues to enhance situational awareness and to localize threats was also discussed.

A novel tactile display device was described by Rupert and colleagues (paper #20) and used to convey attitude information to pilots wearing the device (an array of tactile stimulators wrapped mainly around the torso). With the device, precision in maintaining fixed bank and roll could be maintained after some practice. The formatting of tactile displays to convey basic flight information was also discussed.

5.5 Demonstrations

Several applications were demonstrated on videotape. These included telesurgery, medical

training, architectural design, and voice input/output systems.

6. CONCLUSIONS

No systems using virtual interfaces are yet fielded. With few exceptions (paper #8), system level development also is not yet underway among NATO laboratories reporting here. Conceptual designs and partial implementations, however, are easily found for applications in medicine (paper #3), training (keynote paper and paper #4), and design (paper #9). Impediments to system development currently appear to include three major factors, associated with technology, design, and research. Technological impediments include uncertainty about how best to monitor human performance (papers #4, 5, 6, 10, 11, 12, 13, 14). Design issues concern how to select among numerous design options, in a principled way, by measuring any performance benefits of virtual interfaces (papers #1, 3, 8, 18). Research issues concern how best to predict the perceptual effects of imperfectly rendered or symbolically encoded natural environments (papers #6, 7, 15, 17, 19, 20), how to provide necessary computational resources for virtual displays (papers #9, 18), and perhaps most important for general use, how to eliminate the malaise experienced by a portion of the population using virtual environments (papers #2 and 16).

7. RECOMMENDATIONS

Aerospace medical applications for virtual interfaces are not sufficiently distinct from applications in other domains (e.g. command and control, training and simulation, design) to warrant a completely distinct research and development effort devoted to aeromedicine.

The tools for construction of virtual interfaces derive from multiple disciplines. Research on tools can proceed independently of research on systems. Examples from this meeting include devices for monitoring head, eye and hand movements, and speech production. Other examples include visual, auditory, and tactile displays. Performance monitoring and display technologies have applications in domains broader than virtual interfaces. As a result, other work needs to be done to adapt the tools for use in virtual environments.

The constellation of tools used to implement a demonstration of virtual technology, even if readily available, do not sufficiently constrain the design options. As a result, additional research is needed on concept definition and performance evaluation (against technologies currently used) to demonstrate that virtual solutions to interface problems add unique capability or provide measureable performance benefits.

A RESEARCH AGENDA FOR VIRTUAL ENVIRONMENTS

Professor R.S. Kalawsky
British Aerospace Defence Ltd
Military Aircraft Division
Brough
North Humberside
UK
HU15 1EQ

1. SUMMARY

During recent years a great deal has been written and discussed about Virtual Reality. In fact Virtual Reality has received almost unprecedented press and media coverage. News and views of its capabilities have been made and along with films and amusement games, Virtual Reality has been portrayed to the general public as an experience within a fantasy world. Most people now associate Virtual Reality as a 'new' technology which consists of a helmet mounted display, a glove-like device and a high performance graphics system. They do not realise that Virtual Reality is not a new technology and the aforementioned description of it is only one type of a virtual interface system. Concepts underlying virtual environment systems look set to revolutionise the future aerospace business. With cutbacks in defence spending there is even greater need to employ cost effective measures to improve the efficiency of the business. Applications are likely to range from simulation, cockpit design studies, maintainability assessment, more cost effective training through to complete product visualisation. However, key issues have to be identified and addressed before being developed and applied to a specific task or application.

This paper explains the difference between Virtual Reality and Virtual Environment systems and discusses the requirements of a Virtual Environment System. Moreover, key outstanding research issues are highlighted and recommendations for the way ahead are given.

2. WHAT DO WE MEAN BY VIRTUAL ENVIRONMENTS ?

Virtual Reality, Virtual Environments, Artificial Reality, Cyberspace, and Synthetic Environments are a few of the terms used to describe the same concept. Although Virtual Reality is the term that has become the most popular, a great deal of research has to be undertaken before we can achieve virtual reality. Therefore, the term 'Virtual Environments' seems to be a more appropriate term to use. Ellis[1] (1993) suggests that virtual environment systems are a form of personal simulation system.

Most people associate Virtual Environments with a helmet mounted display, a glove-like device and a high performance graphics system, but such a system is only one type of a virtual interface system.

Today, it is awkward and difficult to define a virtual interface system because as yet there are no clear or consistent definitions to guide us. Many definitions have been proposed, but probably the best abstract description for a virtual interface system is given by Zeltzer[2] (1991).

The definition is based on a model that 'assumes' that any Virtual Environment has three components:

1. A set of models/objects or processes.
2. A means of modifying the states of these models.
3. A range of sensory modalities to allow the participant to experience the virtual environment.

Zeltzer represents these components on a unit cube with vectors relating to autonomy, interaction and presence. (Refer to Figure 1)

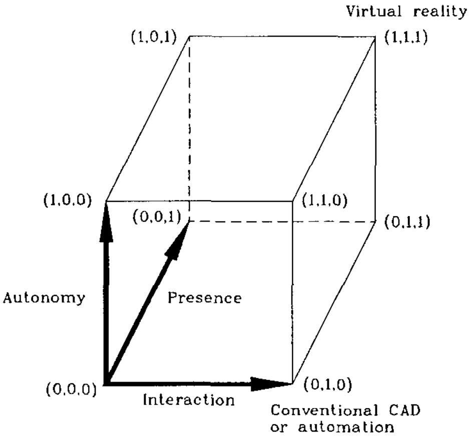


Figure 1 Zeltzer's Autonomy, and Presence Cube

Autonomy refers to a qualitative measure of the virtual object's ability to react to events and stimuli. Where no reaction occurs then the autonomy is 0 whereas for maximum autonomy a value of 1 is assigned. Scaling between 0 and 1 in this context is purely qualitative.

Presented at an AGARD Meeting on 'Virtual Interfaces: Research and Applications', October 1993.

Interaction refers to the degree of access to the parameters or variables of an object. A rating of 0 applies to non real time control of the variables. For example, variables initialised during compilation or at the beginning of execution. A value of 1 is assigned for variables that can be manipulated in real time during program execution. Modern graphics systems allow a very high degree of interaction. However, it is necessary to consider the complexity of the application. A very complex application program may not be able to run in real time.

Presence, or rather the degree of presence provides a crude measure of the fidelity of the sensory input and output channels. The degree of presence has a high dependency on the task requirements - hence the application has a bearing.

At the point (0,0,0) on Zeltzer's cube is represented the very early graphics systems that were programmed in non real time batch mode. These early systems exhibited no interactivity. Examples include graph plotters and chart recorders. Diagonally opposite this point is our aiming point where we have maximum autonomy, interactivity and presence. This is virtual reality. The sensory simulation would be so complete that we would not be able to distinguish the virtual environment from the real world. The point (0,1,0) can be achieved today where the user can control essentially all the variables of an object or model during programm execution. This can be achieved in real time. A point approaching (0,1,1) probably represents the status of virtual environments where we can experience a high degree of interactivity with a reasonable degree of presence. Unfortunately, the degree of automation of the objects in the virtual environment is relatively low. The point (1,0,1) represents the situation where there is a high degree of presence and autonomy but the interactivity is low. An example of this would be a fully autonomous virtual environment where the human becomes a passive observer but is fully immersed in the virtual environment. The only freedom the observer would have is the ability to control their viewpoint. Any change of viewpoint would be oblivious to the objects in the virtual environment.

When the author first attended one of Zeltzer's presentations he was not convinced that the abstract representation of a virtual environment would serve any purpose. However he now finds that when explaining the different categories of virtual environments, Zeltzer's conceptual tool is a very useful aid.

3. OUTSTANDING RESEARCH ISSUES

There are a vast array of issues that relate to a virtual environment system. Whilst our understanding in many areas is quite advanced, our overall understanding of the requirements of a virtual environment system are less clear. Even though we may not necessarily need to achieve virtual reality in the truest sense, we are unable to quantify the requirements of lesser capable systems.

It is tempting to jump on the virtual reality bandwagon and deal only with the technology aspects of the field. However, if the technology is to move forwards then it is be necessary to examine the task before the technology is applied. Only by doing this will it be possible to consider what attributes a virtual environment system brings to the task that cannot be achieved by alternative and lower cost solutions. A business analysis will almost certainly be undertaken which will examine (to a 'first-order' assessment) the technological problems that may be encountered. In many respects the business case will provide the necessary justification for employing a virtual environment system.

3.1 Human Perception in Virtual Environments

Our understanding of human perception and human factors issues regarding virtual environments is still in its infancy. A considerable amount of research in this area is very important because it is needed to focus the development of enabling technologies. Major research areas include:

3.1.1 *Visual perception*

(i) *Spatial resolution* - What display spatial resolution is required for a particular task?

(ii) Field of view is a difficult parameter to specify. However, to achieve an immersive virtual environment a field of view of 100° or more may be required. To achieve a wide field of view a very large optical system is required. The main aim will be to determine what field of view is needed to perform the task effectively.

(iii) *Binocular overlap* - This parameter is related to the total display field of view. To achieve stereo displays a degree of binocular overlap is required. Partial overlapping binocular fields may be used to produce a very wide field of view. However, the amount of binocular overlap is important and must be 'tuned' to suit the application. Perceptual and human performance studies must be undertaken to determine if a partial overlap solution is appropriate.

(iv) *Temporal resolution* - What display update or refresh rate is acceptable for a given task? The higher the update requirement the greater the computational performance will be needed.

(v) *Visual representation* of the virtual environment must be investigated to determine the nature of the scene to be

used for the application. Some applications may require very high fidelity displays whilst other applications may suffice with simplified, cartoon like images. Obviously, there are large differences between these visual representations. How 'real' should the virtual environment appear? The answer must address the spatial and temporal fidelity of the virtual environment. Cost will be a determining factor.

(vi) Is an immersive or desk top system required? (This question can only be answered after consideration of the task, the complexity of the system and the cost.)

3.1.2 *Auditory perception*

(i) In auditory environments the area requiring a great deal more research is the field of 3-D audio localization. Generation of spatialised sound can be achieved with high performance digital signal processors. However, individual differences in pinnae shape can lead to errors when non-personalised head related transfer functions (HRTF) are used. Occasionally a sensation of non-externalization can be experienced. This means that the listener does not perceive the sensation that the sound originates outside the head. Further work is required in characterising HRTF's and determining the causes for lack of externalization in some subjects. Simpler 3-D audio localizer systems do not take account of effects such as reflection and reverberation. These are characteristics of a real environment. Therefore, work must be undertaken to examine the importance of accurate modelling of the acoustical environment. Sound in a real environment undergoes multiple reflections from a range of material types before it reaches the ear. Moreover, sound can be received from a single source via a direct path and many indirect routes. These sound waves combine in the ear to give a very complex waveform. The importance of the secondary reflections and indirect path sound signals must be quantified. If these characteristics have to be modelled it will be important to develop second generation audio localizer systems with an order of magnitude improvement in performance.

(ii) *Improved HRTF* - To achieve an acceptable degree of spatial auditory localization it is necessary to determine the individual's HRTF and use this in the audio localisation system. Ideally, a more generalized solution is required that works for many users and eventually becomes user independent.

(iii) Cues for range and localization. It is known that to determine both range and orientation of the sound signal the type of auditory cue presented to the listener is very important. This is particularly so when first time recognition of sound is required.

(iv) *Externalization* - Many users of spatial sound systems complain that the sound appears to be localized within the head. In other words externalization does not occur. This effect may be a function of the HRTF not being compatible with the listener.

3.1.3 *Haptic/Kinaesthetic Systems*

(i) In comparison to visual and auditory environments, haptic environments are still in their infancy. To maintain a high degree of presence in the virtual environment it is probable that there will have to be direct contact with virtual objects. Discrete approaches are currently being undertaken to stimulate the tactile and kinaesthetic senses. These are largely confined to force reflecting joysticks, hand/arm exoskeletons and tactile feedback gloves. On investigation, the human haptic system is considerably more complex than one realizes. To convey haptic stimulations it is necessary to take account of surface skin and sub-surface physical properties of the tissues.

(ii) The human haptic/kinaesthetic systems need to be characterized and consideration must be given to temporal variations. Manipulation strategies in real world systems should be determined for a range of tasks. Object characteristics such as compliance and roughness must be defined in a way that these parameters can be encapsulated in a CAD/Virtual environment modelling program. To provide computer synthesised haptic responses it will be necessary to develop a computational model of the physical properties of the skin and underlying tissues.

(iii) In order to apply forces to the hand and arm it is necessary to use a form of exoskeleton into which the hand and arm is inserted. The problem of safety must be addressed because forces of the order of 10 Newtons will be applied. There seems to be no alternative to the exoskeleton but to couple haptic and kinaesthetic forces to the operator.

(iv) Work is required in the development of lightweight sensors and actuators to keep the overall mass of the exoskeleton at an acceptable level. The bandwidth and frequency response of a force reflecting system needs to be quantified by careful experimentation. Current tactile stimulation systems are essentially stand alone demonstrations of a field of mechanically activated (or pneumatic) 'points'. Depending on the technology used they either provide small reactive areas (each area covering several square millimetres) or an array of extendable points at a density of 1/2 mm.

(v) A key element to the development of a haptic display system is a complete analysis of the biomechanical properties of the skin.

(vi) *Bandwidth* - To perceive detailed surface texture information it is important to characterize the haptic and kinaesthetic system in terms of bandwidth and dynamic response. If a haptic actuator system is to be built then it must have a bandwidth that exceeds that of the human perception system. A similar requirement exists for force reflective devices, except that the problems of supporting an exoskeleton must be addressed. The actuation system must not only provide the right level of force feedback but it must overcome the mass, inertia and friction of the exoskeleton system.

(vii) *Resolution* - Equally important to the bandwidth of the haptic system is the resolution of the actuator system used to convey the sensation of touch. The spatial resolution and dynamic range are important parameters.

(viii) *Strategies performed by the human with haptic tasks* must be analyzed in a way that allows the actuator technology to be simplified. It is probably impractical to replicate all the cues provided by picking up an object. Therefore it will be necessary to isolate the dominant cues and ensure that these are presented with a sufficient level of fidelity.

4. Performance Metrics

The area of performance metrics is extremely important for determining the effectiveness of a particular virtual environment solution. Without any form of metric it is very difficult to match the human operator to the virtual environment. Moreover, it will be almost impossible to optimise the man machine interface because we have to rely on subjective opinion. The problems of defining suitable performance metrics is not unique to the field of virtual environments. Indeed the whole field of man machine interfacing is desperately calling for a set of standard performance criteria. If a suitable set of metrics were to exist then it would be easy to quantify the benefits that a virtual environment system brings over and above alternative approaches. The author encourages researchers to think very carefully about the advantages of applying a series of metrics to the field of virtual environments. Once a set of metrics has been established then hesitant potential investors may be convinced of the real benefits brought by virtual environment technology.

5. Virtual Environment Technology

(i) *Displays* - Urgent research is required to assist the development of true 1000 x 1000 pixel colour displays. These should be full colour, high update rate and be contained within a small package size of about 25.4 mm square.

High resolution - Future display resolution requirements are likely to approach the limiting resolution of the eye (1 minute of one) with several minutes of arc being a more practical requirement.

Variable resolution - It is well known that the human eye has excellent visual acuity in the region of the fovea. Outside this area the spatial resolution falls off dramatically. It may be possible to develop a display system that is matched to the resolution of the human eye. This will mean using an eye slaved high resolution insert. Eye slaved high resolution patches seem to offer the necessary resolution over relatively small angular subtenses. However, there is a question regarding the performance of the eye tracking technology and the dynamic response of the high resolution patch deflection system. Displays embodying this approach will be expensive.

(ii) *Space tracking technology*

Low phase lag - Without doubt one of the critical areas of space tracking systems (and virtual environments) is the requirement for low phase lag. The phase lag will probably have to be less than 5 mS if the lags are not to affect the performance of the operator. Particular care has to taken when interpreting what is meant by phase lag - as described in Chapter 6.

Resolution requirements for tracking systems probably do not need to exceed 0.1 mm in translation and 0.01° in angular terms. For many applications translation resolution of the order of 1 mm and angular resolution of 0.1° may be quite adequate.

(iii) *Multiple object tracking* - It will be desirable to track multiple objects within a virtual environment. For example - the user's head, and possibly both hands. With most current tracking systems the effective update rate of each tracked object is divided by the number of tracking sensors used. This reduction in update rate is due to the synchronisation or multiplexing of trackers in the system. Unfortunately, this is a consequence of the technology used in the tracking system. A better method of tracking multiple objects is required that does not use multiplexed sensors. Moreover, if the whole body is to be tracked in terms of limb position then this amounts to a considerable number of sensors. Apart from the update problems, the large number of cables connecting the sensors to the tracking electronics becomes a significant problem. Ideally, a wireless tracking system should be used. In the future, image processing systems may be able to determine the position of multiple objects without the need to cable up the participant. However, this will demand considerable processing performance and high resolution imaging sensors.

(iv) *Image generators* - Virtual environments place severe timing constraints on image generation systems. Whilst very high performance can undoubtedly be achieved there is a concern that the architectures of these systems do not lend themselves to the demanding performance required. As described in Chapter 6 the key parameter is the

system throughput time. This figure must be considerably better then current values of 40 mS -100 mS. Apart from designing graphic system architecture to suit the virtual environment application, benefits can also be obtained by employing predictive filtering techniques. Different algorithms must be studied to determine if they offer any real advantage.

Low latency architectures - Current graphics platforms may need to be redesigned with low latency architectures in mind. This requirement derives from the need to couple head tracking systems to the later stages of the graphics system. It is tempting to employ the standard RS232 interface of the graphics system for the space tracking system. Unfortunately, this interface is not usually designed for real time applications. As a consequence, attempts to send large amounts of high speed data through this interface results in an unacceptable interrupt load on the host processor. This means that more time is spent servicing the interrupt than in dealing with graphics drawing operations.

Update rate - The question of update rate is an interesting one. At the moment the computer graphics industry is concerned with increasing the spatial resolution of a display in preference to display update rate. However, for a virtual environment application this may be the complete opposite of what is required. Spatial resolution could be secondary to display update rate. Urgent research is required to determine whether high frame rate displays should be used in favour of high resolution displays. One factor in favour of the high frame display is the limitation in display resolution of current helmet mounted displays. There seems to be little point in wasting computational effort when the display device cannot resolve the fine detail.

Motion prediction - There is some merit in being able to use motion prediction methods to compensate for inherent system lags. Provided the motion of an object can be expressed by means of a motion equation, it is possible that previous motion data can be used to predict where the object will be during the next few iterations. Parameters such as velocity and acceleration profiles are used in the prediction process. In the case of the user's head it will be necessary to determine the dynamics of the human head. Kalman filters could be used to predict where the object or head would be during the next few frames.

6. Virtual Environment Software Engineering

(i) *Visual programming languages* - To build the synthetic environment from a collection of library routines, the majority of software tools for virtual environment applications rely on a competent 'C' programmer being available. In contrast to this the VPL RB2 virtual environment programming tools rely on visual programming techniques. These visual programming techniques allow people, with fairly minimal computer literacy, to create and maintain a virtual environment. With these tools it is possible to create a fully interactive virtual environment without writing any software. The visual programmer constructs the virtual environment by linking objects in the virtual environment with behaviourial constructs. These are represented on screen by icons and simple 'wiring diagrams'. Whilst the highest performance virtual environments will be programmed at a basic level ,the use of a visual programming language will be of great benefit to the person interested in rapid prototyping. As computer graphics systems become more powerful, the performance difference between visual programming languages and conventional programming techniques will converge. As virtual environments become larger, then visual programming techniques may result in significant cost savings that far out-weigh conventional approaches.

(ii) *Database standards* - All virtual environment systems rely on an underlying database standard on which to represent the objects of the environment. In some ways the database is rather like a CAD type database standard. (Some virtual environment software packages are actually based on well known CAD standards.) However, a virtual environment system will generally require considerably more data to describe the environment. Not only is it necessary to describe the geometrical and spatial relationships of objects, but other parameters such as behaviour must be specified. This includes responses to external events or stimuli such as collisions and also includes mass and feel. To date there are no standards in this area. A virtual environment standard is an obvious requirement.

(iii) *Virtual environment modelling*. The whole area of modelling for virtual environments needs attention. At the moment there are no standards developing and there is a danger that future virtual environment systems will have to support multiple standards. If some measure of standardization does not come soon then organisations will have invested effort in their chosen standard and will be reluctant to move to another standard. With a virtual environment it will be necessary to store additional attribute information about an object such as texture (feel), weight, compliance and so on. Therefore, we have an opportunity to develop an open standard that can be used by everyone.

(iv) *Multiple participants* - In order to create multiple participant virtual environments, it will be necessary to develop communication protocols so that consistent databases can be maintained for each user. This means that if one participant moves an object in his virtual environment, then the corresponding object in another participant's environment is updated accordingly. The problems of networking in database systems should be

reasonably well understood. However, some work will be required to ensure that efficient protocols are developed that allow real time operation.

(v) *Use of virtual environments inside the virtual environment systems* - The high level of interactivity within a virtual environment is one of the strengths of the technology. However, this interactivity will only be of value if the design work that is undertaken within the virtual environment can be used outside the virtual environment.

7. PHILOSOPHICAL REFLECTIONS

It is easy to become excited by virtual environments and the potential they offer. However, it is very important to resist this initial burst of enthusiasm and direct one's attention to the task of determining what the key issues of the system should be. It will be necessary to address the nature of the user interface and to understand the system requirements. Only when this has been undertaken should consideration be given to the type of technology that should be employed. Care must also be taken to address the human factors issues inevitably associated with complex man machine interfaces.

An equally important issue that must be addressed along with the human factors and the associated engineering, is a thorough business analysis. Nearly all ventures in high technology systems will fail unless the business issues have been properly addressed. From the author's perspective there are many people who having heard of the term Virtual Reality believe that the subject is all about helmet mounted displays and glove like devices. However, virtual reality or virtual environments is much more than a helmet mounted display and glove device. The business decision makers must be made to understand the wider issues of virtual environments. They must realise that a virtual environment is a synthetic computer generated representation of a physical system. A representation that allows a user to interact with the synthetic environment as if it were real. One of the distinct advantages being that the user is not bounded by limitations presented by the real world. For instance, virtual environments could be used to prototype a product during the early part of its life cycle. The interactivity of a virtual environment would allow the user to explore alternative configurations before the product is manufactured. This approach means that design and development risks could be removed early in the manufacturing life cycle. In many respects the world is already moving towards rapid prototyping systems or synthetic design environments. The benefits of these systems are already established. A virtual environment system addresses the totality of such design and rapid prototyping systems by allowing the user to achieve a higher level of interactivity than can be afforded by computer aided design (CAD) systems. It would be wrong to suggest that every prototyping system will require total immersion in the virtual environment. Some design tasks may be better served by a traditional CAD system but during the latter stages of design a more immersive system may be required. Therefore, a key requirement is the ability to move between these different prototyping systems by providing the designer with the right level of immersion for the task. Ideally the transition between the different prototyping stages would be seamless and extend into the manufacturing process. The concept of assessing ease of manufacture and ease of assembly is extremely exciting. This could be further extended into customer product training whist the product is being manufactured. Manufacturing processes based on a virtual environment could revolutionize the way we design and manufacture things in the future.

8. RECOMMENDATIONS FOR THE WAY AHEAD

There is little doubt that current generation virtual environment peripherals are limited in terms of resolution. However, by conducting research into the human factors requirements it will be possible to match the technology to the human interface. The affordable high resolution full colour helmet mounted display is already on its way and so to are the high performance computer systems. Advances in the other technologies such as tracking systems and haptic/kinaesthetic feedback display systems are moving at a slightly slower pace. As people recognize the importance of virtual environments then improvements will be made. From a virtual environment scientist's point of view it will be necessary to provide human factor's guide lines so that the technology may be appropriately developed. Would be developers of the technology (including software) are advised to consider the standardization of interfaces. This will make it easier to take advantage of improved technology as it emerges. It is hoped that this paper will act as a baseline of knowledge which we can all build up our understanding of the next generation human to machine interface.

9. A STRATEGY FOR FUTURE RESEARCH

The aerospace community is well placed to retain a leading position in the field of virtual environments providing that a coordinated research effort is maintained. Rather than undertake ad-hoc research without any clear objectives it is necessary to agree a research agenda with identifiable objectives and deliverables against clear business drivers.

10. REFERENCES

1. Ellis S.R., (1993) from forward in Kalawsky: The Science of Virtual Reality and Virtual Environments, Addison-Wesley, 1993.

2. Zeltzer D., (1992) *Autonomy, Interaction, and Presence*, Presence, Vol 1, Number 1

3. *Research Directions in Virtual Environments*, Computer Graphics, Volume 26, Number 3, August 1992

4. Kalawsky R.S. *The Science of Virtual Reality and Virtual Environments*, Addison-Wesley, 1993

11. ACKNOWLEDGEMENTS

This paper is based on an extract from a book published by Addison-Wesley Publishers in 1993. Kalawsky: The Science of Virtual Reality and Virtual Environments Addison-Wesley Publishers Ltd © 1993. Finally, kind permission to publish this article by British Aerospace is gratefully acknowledged.

TASK-SPECIFIC USABILITY REQUIREMENTS FOR VIRTUAL INFORMATION ENVIRONMENTS: INTERFACE DESIGN AND DATA REPRESENTATION FOR HUMAN OPERATORS OF COMPLEX MEDICAL SYSTEMS

Michael S. Nilan, Ph.D.
4-206 Center for Science & Technology
School of Information Studies
Syracuse University
Syracuse, NY 13244-4100 USA
Internet: nilan@cat.syr.edu

SUMMARY

The National Research Council has identified "usability" as one of two major requirements for coherent development of computer and information systems over the next ten years [Ref 7]. The use of multisensory virtual environment technology to display and provide access to system functions and data relevant to large-scale, complex, potentially volatile medical tasks (e.g., telepresence surgery) increases the (already critical) need for unobtrusive, transparent interface designs and data representations. Unfortunately, the medical community must take responsibility for providing requirements specifications to the computer industry or else be forced to adapt to existing technical constraints [Ref 10].

Recent research in interface design and data organization/ representation for two dimensional computer applications indicates that dynamic representations of the specific task or problem that the human operator is performing is very effective [Ref 8]. Employing a task-specific, "user-based" methodology, steps in the task resolution are organized into a dynamic model of the task. Linked to this model are the functional system requirements and information/data need requirements divided into specific content requirements, display requirements (including spatial organization), and system help requirements. The resultant model is readily interpretable by system designers and in addition, provides them with specific task-related system evaluation criteria. Usability advantages of dynamic task representations include: minimal system/application training requirements for operators; and coherent, comprehensible and uncluttered sensory field organization of system functions, relevant data and help information. Because of its ability to provide specific task-related requirements to system designers, this methodological approach will insure maximum usability of high performance computing (including virtual reality technology) for critical medical applications.

1. INTRODUCTION - USABILITY CONCERNS

"It is becoming increasingly clear that the comfort of a good fit between man and machine is largely absent from the technology of the information age."

- John Sedgwick, **The Atlantic Monthly**, March 1993, p. 96

Everyone is becoming anxious to solve the usability problem, from popular writers like John Sedgwick [Ref 15] representing users in general to national policy groups like the National Research Council [Ref 7] representing the computer industry, federal policy makers, and academia. There are several conditions that have generated this interest in usability including:

- a trend towards distributed computing along with the accompanying increase in complexity for users;
- a general shift from a manufacturing economy to a service economy, i.e., a shift from a product orientation to a service orientation over the last twenty years or so and, in the computer industry, over the last couple of years;
- the current economic recession, particularly in computer related industries, e.g., the newest hardware platforms and software updates haven't sold very well;
- management in user organizations' concerns with the "hidden" cost of training associated with new applications, application updates or new workers; and
- very widespread frustration of users in general with the lack of simple coherency in system design, particularly across applications.

Presented at an AGARD Meeting on 'Virtual Interfaces: Research and Applications', October 1993.

The introduction of the IBM Personal Computer just over ten years ago, in addition to stimulating the information explosion, was a major catalyst in the spread of computerized systems from laboratories and data processing departments to virtually every aspect of human activity. Until very recently, the concern for usability was primarily the concern of vendor marketing departments while the "real" system designers focused on smaller, faster, flashier, more gimmicks, etc. based upon the technology (e.g., the perplexing profusion of graphic icons in Microsoft Word, version 5.1 and just about all virtual reality applications). In essence, the range of capabilities of systems have not matched well with the range of user needs and system features have been represented to users in a confusing variety of cryptic forms. This is not to say that designers weren't interested in usability but rather that it was not as high on their agendas as it has been on users' agendas. And yet, the aphorism about hardware being ten years ahead of the software persists.

Concurrent with this shift within systems development organizations towards more usable systems, the users were learning a few lessons as well. They have learned, for example, that their most important investment is in their data and being able to easily employ that data to solve problems, make decisions and plan rather than investing in the newest, fastest, highest resolution hardware. They are also getting a good sense of how much time and energy needs to be invested in learning/training to get existing systems to do even the rudimentary data manipulations that the systems are capable of, and these systems *still* don't do what the users need.

In spite of the incredible things that computerized systems can do, the feeling users are getting is that most of these systems are solutions running around looking for problems. Users already have problems and those problems are not adequately reflected in existing systems. While there are some very notable exceptions to this, e.g., Lotus 1-2-3, for the most part, the American free enterprise notion of inventing something and then marketing it has been a serious impediment to addressing users' needs, a problem originating at the management level of system design organizations. After all, if the market was buying the systems, why change? Consequently, usability in existing systems is quite poor. If this is the case for so-called "stand alone" applications, it is doubly so for distributed high performance computing and communications (HPCC) technology that might be able to facilitate complex medical problems.

2. SYSTEMS DESIGN: IN THE BEGINNING...

A major source of assumptions and insight into computer system design comes from the seminal work of Herbert Simon and Alan Newell. In **The Sciences of the Artificial** [Ref 16], Simon established an approach to system design based upon simulation of human cognitive processes (i.e., creating systems which demonstrate the ability to arrive at a functionally equivalent solution to a problem or task) as a means by which designers can learn about designing systems and, at the same time, learn about human cognition. His justification for this "isomorphism" approach is that there are two sources of insight into human cognition, the internal (i.e., what actually goes on when people think) and the external (i.e., watching the actions people use to solve problems, etc. and developing functional simulations of the process and outcome). Simon argued that the internal source of insight is not available to designers but that is all right because the external is just as good. At least two generations of system designers (i.e., computer scientists, computer engineers, cognitive psychologists, programmers, etc.) have followed this assumption in their approach to design through their approaches to the user interface and data representation. Note that this is the era that led up to the current situation where usability has become much more essential to effective system design and even essential to the economic prosperity of the United States [Ref 7].

One way to interpret Simon's argument is that somehow, technology stands apart or is different from human behavior. I would propose a different picture of technology, i.e., that rather than being something separate from human cognition, I would argue that all technology is a derivative of human cognition. The etymological origin of the word "technology" stems from "technique" and even further back to the Greek "technos," both of which essentially refer to the process by which a problem is solved or something is accomplished. In fact, virtually *all* technological applications are an extension of human cognitive, sensory and motor capabilities [Ref 9]. Further, one of the most serious problems with systems development right now can be explained by a process that builds new technology on top of old technology without effectively checking back with the human problem that stimulated the application in the first place. A vicious circle is established where technological applications are supplying insight into design rather than the original human needs. This is particularly troublesome since we weren't very good at understanding

user needs when we designed the old technology. In this sense, it would seem that Simon's argument is a bit tautological and, with regards the usability problem, is not likely to provide much insight into users' needs. In other words, I disagree with Simon that the internal is inaccessible and I feel that the external is not sufficient for usability concerns. How can technology NOT be inherently tied to human perceptions of the problem being addressed by the system? One of the other assumptions of the user-based approach is that the ideal model for human-computer interaction is human to human interpersonal communication. This means that one of the essential "places" to search for insight into system design is internal; exactly opposite from the strategy espoused by Simon.

One of the costs to usability that has resulted from the adoption of Simon's logic is that systems have been designed, presented to the user, and the user must adapt his/her behavior to the system, i.e., the user must become more like the system in order to effectively employ it. The user-based approach argues that systems must become more like users [Ref 3] and that this is not only do-able, it is imperative.

Using a database application as an example, early database management systems (DBMS) were highly "structured" in that the user (after spending a lot of time learning the application) could do only a very few things that had been programmed into the system. The Simon-driven response to user dissatisfaction has been to create (relatively) "unstructured" DBMS systems that are supposed to allow the user to do anything the programmers could dream up plus anything the users could articulate indicating what they "wanted" (see [Ref 3]). Users are no more experts in knowledge acquisition than are system designers. As a result, the user is overwhelmed by the variety/complexity and has no idea how to proceed to make effective use of the system functionalities (even if the representation was comprehensive and somehow comprehensible through training). What is really needed is a design approach that emphasizes functions known to be useful to users which are then represented in a semi-structured manner so that the user has guidance from other users who have solved the same problem, made the same decision, etc., rather than menus or ranges established by designers and programmers. What is not clear in practice or in the literature is how designers might do this or whether users might more productively specify what they need because designers are obviously not doing this.

In practice, the focus on usability is often placed on the user interface and this remains a very vexing problem in system design:

> "Not only is there disagreement about how to arrive at a good user interface design, it is not even clear who, on the development team, should be responsible for this task. In short, we don't know who or what kind of knowledge is most advantageous in producing good interface design. The responsibility for the user interface has been relegated to a variety of people with a variety of backgrounds (Bailey, [Ref 1]).

There are other aspects of system design that effect usability, of course, including the logic by which the user cognitively organizes the various functions that the system offers (including which functionalities of solving the particular problem have been incorporated into the system), the way that data are organized in the system (both "help" data or documentation as well as data that are being manipulated by the system), and data representation or the form that the data are given to the user. The argument here is that we need to adopt a strategy for understanding users that is quite different from Simon's if we are to coherently address usability. The reader should note that I am not quarreling with all of Simon's approach but rather those aspects that are not effective for usability concerns.

3. MEDICAL COMMUNITY NEEDS

The medical community is notably behind other scientific counterparts in the deployment of computing technology. A number of reasons contribute to this lag in adoption including:

- the relative lack of specialized systems appropriate for medical practitioner and administrative problems beyond office automation;
- the already incredible intellectual demands on training physicians that leaves little or no room for specialized information technology training;
- the expense of existing HPCC medical systems (e.g., CAT, NMRI);
- the high marketing pressure on and resultant confusion of physicians and hospital administrators for a wide variety of computerized (and non-computerized) systems;

- the potential complexity of many medical task situations (both in terms of specific activities within a particular task as well as in terms of the complexities within an activity); and

- the relatively large investment in existing systems (so-called "legacy" systems).

Not only are system designers not effectively representing the user in generalized applications, but they are philosophically and methodologically unable to understand the unique needs of the medical community. For example, from the technological perspective, the imaging needs of physicians and surgeons do not always (or even very often) require the very high resolutions that might be useful in other scientific contexts. The surgeon for example, needs to "see" what s/he is doing and what is behind the externally visible surface. A resolution of 640 X 480 would be quite sufficient if, what was underneath the tissue that is being cut with a scalpel is visible as well. The issue here is not one of resolution but making visible "hidden" aspects of the problem at hand. Processing speed of the image display may be an issue, but resolution is not (in the vast majority of cases). For the majority of surgical procedures that are carried out, the precision of the scalpel plus or minus 2 mm is sufficient. The "real" problem is revealing hidden layers accurately. Another example might be a physician examining MRI data. The problem is differentiating one tissue type from another (e.g., healthy from diseased tissue). Again, high resolution or three-dimensional views are not likely to address this visualization problem. Unfortunately, none of the existing work on visualization has addressed this specific combination of needs.

On the administrative side, particularly in mobile military health care communities, maintaining patient record systems which include text as well as images and sound is an extremely difficult distributed problem. These records may be needed by a surgeon in an operating theater and at the same time by a specialist thousands of miles away. While the technical capability might be available, because of the relative crudeness of user interfaces, task-oriented data management systems, and ad hoc data representation, it is currently not feasible.

So, in spite of the relatively impressive gains in computing power, speed, bandwidth, imaging, etc., the "fit" for a reasonable deployment of HPCC information and computing systems in medicine remains elusive. There are two very compelling reasons however, that make medicine a perfect market segment to force the issue of usability:

- the medical community lags behind other scientific and engineering professions in the use of high performance computing and communications; and

- the medical community represents a VERY large market (particularly of late due to the potential of HPCC to contribute to reducing the costs of health care reform).

In a real sense however, most high performance computing and communications, including virtual environment technologies, represent "solutions looking for a problem" [Ref 10]. What this paper will argue however, is that if the concern is usability (as defined above), the idea is NOT to fit problems to available technology but to fit technology (and/or to develop technology) to address problems as those problems are understood by the users. To see how users understand problems, we need to look at knowledge acquisition and representation.

4. KNOWLEDGE ACQUISITION AND REPRESENTATION

Basically, what is needed in system design is a way of interacting with users that generates lists of functions that users employ to solve their problems, etc. and a way to represent those functions to system designers so that the resulting system is inherently (or with minimal learning required to be) understandable to users. The most important "place" this must be done is where users cognitively interact with the system, i.e., at the user interface and where data is represented to the user (or the user employs an internal data representation to search for data relevant to the task at hand).

Recently, there have been some interesting developments in conceptualizing approaches for understanding user needs (e.g., [Ref 18]). Among the more innovative approaches to understanding users, I would include Lucy Suchman's [Ref 17] efforts at Xerox to try ethnographic methods to find out what users need. She is immersing herself in actual problem or decision contexts (what Dennis Wixon [Ref 19] calls "contextual design") and observing the dynamics. Donald Norman, who is currently working with Apple, employs his "user-centered" design [Ref 14] via "cognitive engineering" so that he can make better design decisions. The Association for Computing

Machinery's Special Interest Group on Computer and Human Interaction (ACM/SIGCHI) has begun to discuss alternative approaches to figuring out exactly what it is that users want (e.g., [Ref 6] discussing ethnographic versus experimental psychological approaches). However, there is no agreement on "what" should be looked at in human behavior to insure usability nor is there an agreement on "how," i.e., methods.

In terms of "what" to look at, the relationship between human beings and technology suggests that we look at how people perceive that they solve problems, make decisions, plan, etc. and see if there are any patterns across users; we start with the problem rather than with what the technology can or might be able to do. We cannot literally get inside peoples' heads, but we can develop very detailed and valid pictures of how people perceive their problem solving processes. We know we can do this because people teach each other through language (either in direct conversations or vicariously via books, articles, training manuals, etc.). Fortunately, there is a growing body of literature (e.g., [Ref 4], [Ref 8]) that indicates:

- people experience a problem or decision as a sequence of actions or steps over time;
- although there are some differences in the amount of detail between novices and experts in how they perceive a particular problem, there are also distinct patterns common to both, i.e., certain actions or steps that are taken in the same time order; and
- there are also patterns in the language that users employ to refer to actions or steps (because they are trying to communicate).

So, for the knowledge acquisition aspect of system design, the user-based approach has adopted observation techniques and strategies from clinical psychology, ethnomethodology and communication science. The resulting frame-based methodology (e.g., see [Ref 2]) has users describe their view of the problem solving process in the order that the actions or steps occurred to them (this can be done with recall techniques or in real-time). The researcher represents these "action objects" to the user as a sequence (on three by five cards for example) that is intended to be isomorphic with what the user actually did (or is doing) in solving the problem. This idiosyncratic model of the steps in the problem solving process is then used as a dynamic mnemonic device to elicit further details about what the user was thinking, what information the user needed at that point in time, what tools were used at that point in time, what the user was trying to accomplish, etc. This approach has several advantages:

- the user-based researcher is asking the user to communicate his/her actual experience with a problem (i.e., one representation of the internal experience) rather than a hypothetical or experimental problem;
- because this structure is used to probe for detail in the user's internal experience, it allows for time-specific details to be probed in more detail; and
- all elaborations on the cognitive experience of the user are linked to this dynamic mnemonic structure (the utility of this analytically will be discussed below).

The resulting knowledge structure has both temporal and spatial structural features. The arrangement of the steps in the problem in a time order should be able to help system designers know "when" the user will need system functions or help or specific kinds of data. The "what" the user needs is a spatial representation of the cognitive activities of the user's needs at that point in time. Note that the spatial features can be arranged according to their importance or their frequency of use across users, etc.

One of the most common complaints that traditional knowledge engineers have at this point is that they feel that there is too much variance in the ways that human beings perceive their environments leading to the conclusion that we cannot use this kind of perceptual data to create knowledge representations suitable for system design. However, if there really is chaotic variance, language for communication should be impossible (and it obviously isn't) and technological applications would have to be different for each person (and they don't). So, the user-based approach employs users' dynamic descriptions of the way that they perceive their own problem solving, deciding, planning in a task specific context.

There are some powerful advantages as a "side effect" of this source of insight into user needs. For example, when people talk, they have to talk about one thing at a time and they usually start at the beginning and end at the end. We know that conscious attention is serial, i.e., one thing at a time. We also know that people experience existence (and therefore problem solutions, decisions, etc.) as changes over time. The isomorphism here between cognitive functions, experience and communication is a powerful source of

validity. Another advantage of this approach is that users tend to employ language that is inherently meaningful to them and the person they are talking to (i.e., the user-based knowledge engineer). The communication capability is a significant usability consideration that is passes, along with the knowledge structure, to the system designer (as discussed below).

Although there are a variety of specific methods suitable for in-person interviews, self-reporting, etc., there are also a number of semi-automated methods that could be employed as long as the user-based conceptual approach is taken as a guideline. For example, Maes & Kozierok at the MIT Media Lab [Ref 5] are working on neural network-based "intelligent agents" which collect data about how a particular user does things on the system, detects patterns in the user's behavior, and eventually does those things for the user. By employing such an agent at a "higher" level in the system, designers could learn about patterns *across* users in a real-time fashion. Another example is "CommTool" being developed for the New York State Center for Advanced Technology in Computer Applications and Software Engineering [Ref 13]. This is a system-level software tool that allows users to communicate with system designers from the rapid prototyping stage through the implementation stage and into the maintenance stage (referring to the software life cycle). In addition, CommTool also provides analytic guidance for directing the user feedback to the appropriate system person (i.e., information/data providers, system maintenance people, system designers and analysts, managers, etc.).

The way this type of knowledge is represented for system designers is identical with the analytic procedures used by researchers to interpret results across users (see [Ref 11] for a detailed example with a desktop publishing application). An "action by cognition" matrix is created where the steps in the problem solving process (in time order) that all users agreed upon constitute the horizontal dimension. This is the basic interface structure. All user access to system functions, to help information, to data created in the process of using the application, etc. are via this "action" dimension of the matrix. For each one of these steps, all of the activities across users in between and including the agreed upon step are set up as a secondary action level. At this level, an expert would see everything s/he needs to do at a particular point in the problem solving process and a novice would immediately get an idea of what kinds of things are possible. Linked (in a hypertext sense) to this secondary action level, are all the system functions necessary for completing that particular step in the process. Included in system functions are network communications links, specialized display requirements, etc. Also linked are all help files so that a novice could get help directly oriented to the particular point in the problem solving process that s/he was at that point in time. Finally, supporting data either commonly used in the problem solving process or that is produced as a byproduct of the process are also managed from this secondary level. The range of system functions, help messages, and data that are linked to each action step represent the system assistance in the "cognitive" aspects of the problem solving process. While this description is obviously an abstract characterization and some of the nitty gritty details are left out (see [Ref 11] for a more detailed description), the "action by cognition" model is really a cognitive communication model because:

- first it is used for establishing communications between the user-based knowledge engineer and the users;
- second, a derivation of the model is used to a communicate the users' needs to the system designer;
- then it is employed to represent the system to the user in the interface; and
- finally, the model can be used as a system evaluation tool with a variety of task-based evaluation criteria beyond the normal "time on task" or "percentage correct" measures (which actually measure user performance, not system performance).

Regarding virtual reality technologies and their incorporation into more complex task situations, a conceptual proof of concept of a "system design space" was implemented for the U. S. Air Force Office of Scientific Research [Ref 12]. The purpose of this project was to illustrate how even extremely complex, distributed, high performance computing and communications architectures and performance requirements could be linked to complex user-based knowledge structures for specific tasks. The demonstration of the capability was implemented on virtual reality technology at Rome Labs (Griffiss Air Force Base) in New York State. While complexity *per se* is not the point of employing task-specific usability requirements as the foundation for system design, complexity in the form of detailed process, associated information needs, associated (both computerized and non-computerized) tools and sub-processes, data management across

a distributed organizational network, etc. can ALL be coordinated in this manner. In fact, this is the only manner in the literature where all system design considerations can be represented in the same "design space."

5. CONCLUSIONS

The argument of this paper is that the user-based approach goes beyond the existing approaches to system design in the following ways:

- the role of the user is seen as conceptually central to the design process because human-computer interaction itself is seen as a basic human-to-human process;
- instead of asking the users what they want or need and instead of observing users' external behaviors and then inferring what is going on inside their heads, users' knowledge/experience is an integral part of the user-based approach; and
- users' descriptions of their perceptions of problem solving, decision making and planning processes are used to create dynamic knowledge structures that are used to stimulate the users' memory, to facilitate the communication between knowledge engineers and system designers, to provide a knowledge representation structure for interface design and data organization, and is used for subsequent system evaluation.

The user-based approach to knowledge acquisition, knowledge representation and (recently) to information system evaluation, is a relatively new approach. It has, however, been tested, evaluated, and found to be extremely useful in addressing the usability problem in system design. The reader should be aware that the user-based approach represents another source of insight into the design process. Existing efforts by cognitive psychologists, computer engineers, software engineers, etc. need to be continued, but any system design team should have at least one user-based knowledge engineer to insure that users are adequately represented in the design process.

For the medical community, this paper argues that historical forces (e.g., increasing demands for usability, health care reform), economic forces (e.g., recession in the computer industry, potential market that the medical community represents), the increasing gap between technological capabilities and the medical community's exploitation of that technology, and, most importantly, the needs of the medical community itself can be well served by developing user-based descriptions of their needs for presentation to system designers in the computer industry. Coalitions among health care organizations (both public and private) as well as coalitions among categories of health care professionals (e.g., laproscopic surgeons, health care maintenance organization administrators, military health care administrators, etc.) can reduce the cost and dramatically increase the effectiveness of HPCC and virtual environments in solving their information and communication needs. While virtual environments represent great potential for addressing complexity issues and, as an emerging technology have their own technical constraints, the issue of usability is fundamental to the future of all computing technologies. The medical community can take the lead in usability improvements, help develop HPCC to address its own problems and at the same time facilitate improvements in usability for other user communities.

ACKNOWLEDGMENTS

The author would like to acknowledge the assistance of the School of Information Studies at Syracuse University, the New York State Center for Advanced Technology in Computer Applications and Software Engineering, the Alex Nason Foundation of New York City, and the U. S. Air Force Office of Scientific Research for their support of the various research projects reported on in this paper. The author would also like to thank the following Ph.D. students, without whom, these projects would have been impossible: Joanne Silverstein, R. David Lankes, Bob Travica, and Bhaskaran Balakrishnan.

REFERENCES

1. Bailey, G. (1993). Iterative Methodology and Designer Training in Human-Computer Interface Design. In the **Proceedings of INTERCHI, 1993** (Amsterdam, The Netherlands, 24 April-29 April, 1993) New York: ACM, pp. 198-205.

2. Dervin, B. (1983). An Overview of Sense-Making Research: Concepts, Methods, and Results to Date. Paper presented at the International Communication Association Annual meeting (May 26 - May 30, 1983, Dallas).

3. Dervin, B. & Nilan, M. S. (1986). Information Needs and Uses. In M. E. Williams (Ed.), **Annual Review of Information Science & Technology**, 21, pp. 3-33.

4. Katzer, J. & Fletcher, P. T. (1992). The Information Environment of Managers. In M. E. Williams (Ed), **Annual Review of Science and Technology**, 27, pp. 227-263.

5. Maes, P. & Kozierok, R. (in press). Learning Interface Agents. **Proceedings of American Association for Artificial Intelligence '93**.

6. Nardi, B., Gilbert, N., Mantei, M., & McCarthy, J. (1993). Mixing oil and water? Ethnography versus Experimental Psychology in the study of computer-mediated communication. In the **Proceedings of INTERCHI, 1993** (Amsterdam, The Netherlands, 24 April-29 April, 1993) New York: ACM, pp. 3-6.

7. National Research Council (1992). **Computing the Future: A Broader Agenda for Computer Science and Engineering.** Washington, D. C.: National Academy Press.

8. Nilan, M. S. (1992). Cognitive Space: Using Virtual Reality for Large Information Resource Management Problems. **Journal of Communication**, 42:4, pp. 115-135.

9. Nilan, M.S. (1993). Speculations on the Impact of Global Electronic Networks on Human Cognition and Human Organization. **Electronic Networking**, (in press). Copies available from the author.

10. Nilan, M. S., Silverstein, J. L. & Lankes, R. D. (1993) Whose Reality? The VR Technology Agenda in Medicine. **AI Expert Virtual Reality '93: Special Report**, pp. 33-37.

11. Nilan, M. S. (in preparation). User-Based Information/ Communication Systems: Using Sense-Making to Create a Desktop Publishing Help and Training System. In B. Dervin (Ed.), **Methodology Between the Cracks: Theory and Method for the Qualitative and Quantitative Study of Human Communication**. Copies available from the author.

12. Nilan, M.S. & Lankes, R.D. (1992). User-Based Requirements for Large-Scale Distributed Information Management Systems: Representations for System Designers. Report to the U. S. Air Force Office of Scientific Research.

13. Nilan, M. S. & Travica, B. (in preparation). CommTool: A Tool of User-Based Software Design, Refinement and Evaluation. Copies are available from the author.

14. Norman, D. A. (1986). Cognitive Engineering. In D. A. Norman & S. W. Draper (Eds.), **User Centered System Design: New Perspectives on Human-Computer Interaction**. Hillsdale, NJ: Lawrence Erlbaum Associates, pp. 31-61.

15. Sedgwick, J. (1993). The Complexity Problem. **The Atlantic Monthly**, 27:3, pp. 96-104.

16. Simon, H. A. (1969). **The Sciences of the Artificial**. Cambridge, MA: MIT Press.

17. Suchman, L. (1987). **Plans and Situated Actions: The Problem of Human-Machine Communication**. Cambridge: Cambridge University Press.

18. Winograd, T. & Flores, F. (1986). **Understanding Computers and Cognition: A New Foundation for Design**. Menlo Park, CA: Addison-Wesley.

19. Wixon, D., Holtzblatt, K. & Knox, S. (1990). Contextual Design: An Emergent View of System Design. In **Proceedings of the Conference on Human Factors in Computing Systems** (Seattle, April 1 - April 5, 1990), New York: ACM, pp. 331-336.

PROFILE ANALYSIS OF AFTER-EFFECTS EXPERIENCED DURING EXPOSURE TO SEVERAL VIRTUAL REALITY ENVIRONMENTS

Robert S. Kennedy, PhD
Essex Corporation
1040 Woodcock Road, Suite 227
Orlando, Fl 32803

Marshall B. Jones, PhD
Behavioral Sciences Department
Milton S. Hershey Medical Center
500 University Drive
Hershey, PA 17033

CDR Michael G. Lilienthal
Defense Modeling and Simulation Office
1901 N. Beaureguard Street, #504
Alexandria, VA 22311

Deborah L. Harm, PhD
Code SD-5Johnson Space Center
Houston, TX 77058

SUMMARY

Motion sickness symptoms are an unwanted by-product of exposure to virtual environments. This problem is not new and was reported in the early flight simulators and experiments on ego motions and vection. The cardinal symptom of motion sickness is, of course, vomiting, but this symptom is ordinarily preceded by a variety of other symptoms. In his classic studies of motion sickness conducted before and during World War II, G. R. Wendt introduced a three point scale to score motion sickness beyond a vomit/no-vomit dichotomy. Later, Navy scientists developed a Motion Sickness Questionnaire (MSQ), originally for use in a slowly rotating room. In the last 20 years the MSQ has been used in a series of studies of air, sea, and space sickness. Only recently, however, has it been appreciated that symptom patterns in the MSQ are not uniform but vary with the way sickness is induced. In seasickness, for example, nausea is the most prominent symptom. In Navy simulators, however, the most common symptom is eye strain, especially when cathode ray tubes are employed in the simulation. The latter result was obtained in a survey of over 1,500 pilot exposures. Using this database, Essex scientists conducted a factor analysis of the MSQ. We found that signs and symptoms of motion sickness fell mainly into three clusters: 1) oculomotor disturbance, 2) nausea and related neurovegetative problems, and 3) disorientation, ataxia, and vertigo. We have since rescored the MSQ results obtained in Navy simulators in terms of these three components. We have also compared these and other profiles obtained from three different virtual reality systems to profiles obtained in sea sickness, space sickness, and alcohol intoxication. We will show examples of these various profiles and point out simularities and differences among them which indicate aspects of what might be called "virtual-reality sickness".

Presented at an AGARD Meeting on 'Virtual Interfaces: Research and Applications', October 1993.

1 INTRODUCTION

In many areas of advancing technology, it is not uncommon to find unwanted by-products. These negative consequences can become serious problems if they are not anticipated and resolved early in the systems development process. Motion sickness like symptoms, disequilibrium and other post effects are examples of the problems faced by the training device industry in the past ten years and have been termed simulator sickness (Crosby & Kennedy, 1982).

Simulators, by design, present rearranged and altered perceptual worlds and are a sub class of the newly developing virtual reality (VR) systems. We believe that VR sickness is likely to occur and should be addressed as technologies develop.

Simulator sickness was first reported over 30 years ago in two studies by Havron and Butler (1957) and Miller and Goodson (1960). Since that time, the numbers of studies and reports of simulator sickness have increased at an exponential rate; there was as much published on simulator sickness since 1990 as in all previous years. A simulator sickness program, sponsored by the U.S. Naval Air Systems Command, began in a formal way in 1982 and initially emphasized problem definition. A series of simulators was surveyed and the incidence documented (Kennedy, Lilienthal, Berbaum, Baltzley & McCauley, 1987). Then the U. S. Navy sponsored two workshops attended by persons knowledgeable in visual vestibular interactions. Reports from the workshops, in the form of guidelines and suggestions for research (Kennedy, Berbaum, Lilienthal, Dunlap, Mulligan, & Funaro, 1987) resulted in a field manual (NTSC, Simulator Sickness, 1989) which is currently in use in the U.S. and some NATO countries. Since then, the emphasis has shifted to the identification of the nauseogenic properties of the stimulus, particularly the inertial forces and, to some extent, the visual characteristics of the stimulus.

Crucial to the design of simulators is specification of the equipment parameters that will promote training effectiveness and realism, but also avoid simulator sickness. However, the technological advances which have provided the opportunity for increased fidelity have, in turn, placed greater demands on other tolerances on simulator subsystems (e.g., responses of visual and motion base systems and their interaction). Visual display systems combine diverse methodologies for generating and enhancing visual information, and sometimes through misalignment, failure, or other factors, eyestrain and other symptoms related to motion sickness may be experienced. Yet pilots may be unaware of the source of these difficulties and are therefore sometimes unable to provide enough information for the visual display engineer to identify and correct the problem. Needless to say, standards and specifications to address these problems are also lacking.

As more and more facilities have begun human factors programs to develop virtual environments for training, operational and recreational usage, aftereffects are increasingly being reported in much the same manner as was found in simulator usage. Indeed, a recent issue of Presence (Vol 1, Number 3, Summer 1992) had been devoted to articles which related simulator sickness applications to virtual reality systems and at a recent conference (Virtual Reality Annual International Symposion, 1993) on virtual reality (VR) systems, several papers alluded to the requirement for virtual reality technologists to attend to visual vestibular interactions since they are the likely source of VR sickness. We predict that VR sickness will be sufficiently like other forms of motion sickness and simulator sickness that important

diagnostic information is available by making comparisons of the symptom data. Historically, scientists involved in the experimental study of motion sickness employ motion sickness symptomatology questionnaires (Kennedy, Tolhurst, & Graybiel, 1965) to handle the problem of different symptoms being experienced by individuals. The MSQ reflects the polysymptomatic nature of simulator sickness in that multiple symptoms are taken into account in the diagnostic scoring.

The theory behind scaling motion sickness severity is that vomiting, the cardinal sign of motion sickness, is ordinarily preceded by a combination of symptoms (Lentz & Guedry, 1978; McNally & Stuart, 1942; Money, 1970). Therefore, in order to score motion sickness beyond merely a vomit/no-vomit dichotomy, Wendt (1968) initially employed a three-point continuum scale in a series of studies on motion sickness. This scale was used to assess motion sickness symptomatology, whereby vomiting was rated higher than "nausea without vomiting" which, in turn, was rated higher than discomfort. Navy scientists developed a Motion Sickness Questionnaire (MSQ) consisting of a checklist of symptoms ordinarily associated with motion sickness for use in sea and air sickness studies (Kennedy et al., 1965). These symptoms included: cerebral (e.g., headache), gastrointestinal (e.g., nausea, burping, emesis), psychological (e.g., anxiety, depression, apathy), and other less characteristic indicants of motion sickness such as "fullness of the head." A response was required for each symptom using a rating of "none", "slight", "moderate", or "severe" (or in some cases "yes" or "no"). From this checklist, a diagnostic scoring procedure was applied resulting in a single, five-point symptomatology scale, serving as a global score reflecting overall discomfort. The five point scale was expanded in studies of seasickness conducted by the U. S. Coast Guard, with the cooperation of the U.S. Navy, (Wiker & Pepper, 1978; Wiker, Kennedy, McCauley, & Pepper 1979a, b; Wiker, Pepper, & McCauley, 1981). These scoring techniques are useful in that they permit quantitative analyses and comparisons of motion sickness in different conditions, exposures, and environments. However, a deficiency for the study of simulator sickness is that the single global score does not reveal information about the potentially separable dimensions of simulator sickness and it lacked statistical normalization properties. It was argued that such information could be informative about the nature of simulator sickness and may also serve a diagnostic function; not just about the individual but to signal differences in the equipment factors (e.g., visual distortion; motion characteristics) which may differentially cause the sickness.

2 METHOD

Simulator Sickness Questionnaire (SSQ)

In order to obtain information about separable dimensions of simulator sickness, >1000 Motion Sickness Questionnaires (MSQ) have been factor analyzed (Lane & Kennedy, 1988; Kennedy, Lane, Berbaum & Lilienthal, 1993). The results of that study produced three specific factors and one general factor. The three factors form the basis for three SSQ subscales. These subscales or dimensions appear to operate through different "target" systems in the human to produce undesirable symptoms. Scores on the Nausea (N) subscale are based on the report of symptoms which relate to gastrointestinal distress such as nausea, stomach awareness, salivation, and burping. Scores on the Visuomotor (V) subscale reflect the report of oculomotor-related symptoms such as eyestrain, difficulty focusing, blurred vision, and headache. Scores on the Disorientation (D) subscale are related to vestibular disarrangement such as dizziness and vertigo.

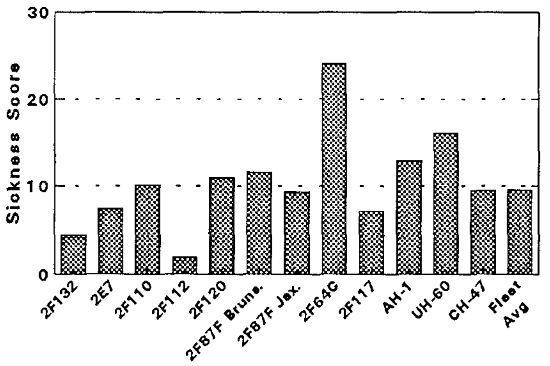
It was also found that the list of symptoms could be abbreviated with little loss in accuracy. Subsequently, a Simulator Sickness Questionnaire (SSQ) was developed based on 16 symptoms only. In addition to the three subscales, an overall Total Severity (TS) score, similar in meaning to the old MSQ score, is obtained. Each SSQ subscale was scaled to have a zero point and a standard deviation of 15. The scoring of the questionnaire is shown in Table 1.

Figure 1 shows total simulator sickness scores for several simulators in the Navy's inventory.

Total Sickness Scores

Simsick Database

Figure 1



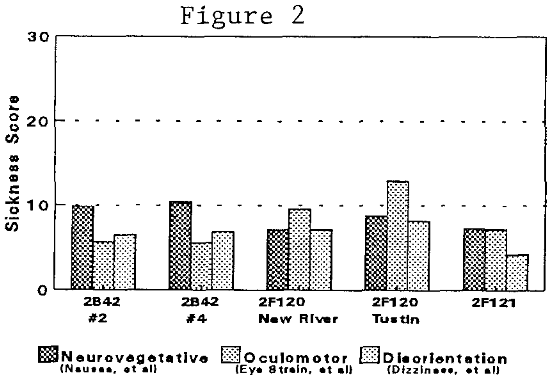
Note that sickness in simulators varies from an average near zero (2F132, a fixed base operational flight trainer for the F/A 18) to an average near 20 (2F64C, a moving base helicopter weapons system trainer for the SH-3). These scores which are used to evaluate the performance of the simulator are presently employed as arithmetic means but ordinarily the incidence of sickness is positively skewed. Therefore, even a simulator with a low score may still place some pilots at risk after leaving the simulator (Kennedy et al., 1987). Therefore, we recommend the use of an additional score to index the safety of a simulator. In our view, anyone with a score higher than 20 (i.e., 1.3 stand deviations) should be warned of his/her condition and not permitted to leave the simulator building unless extreme care is used. Anyone with a score over 15 (i.e., one standard deviation) should contact a flight surgeon or corpsman or be carefully debriefed by an experienced instructor pilot. We also believe that the score attained by the 75th percentile person may be a useful index in this regard. In addition to the total scores, it is possible to use the factor scores as a kind of profile. We find it informative to report factor scores for each of the simulators and to compare them to each other as well as to other forms of motion sickness. We believe that following differential diagnosis of simulator sickness inferences can be made about cause and remediation can be made from these comparisons.

Figure 2 shows the profile score from five helicopter simulators (2B42 [#'s 2 and 4] NAS Milton FL; 2F117 one in MCAS New River NC and one in MCAS Tustin CA & 2F120 in MCAS New River).

Profiles of Simulator Sickness

Helicopter Simulators

Figure 2

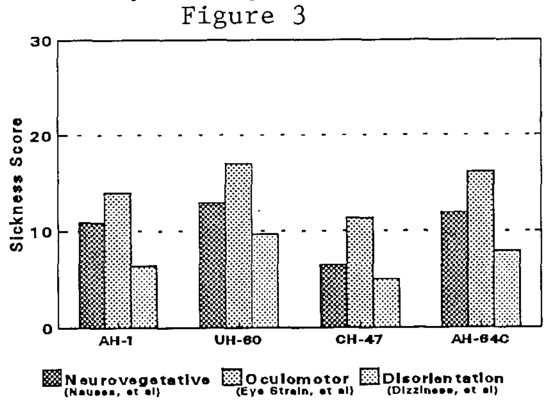


It may be seen that the twin simulators in the same city show mirror images of symptoms as do the twin simulators with the same designator in different cities. The one simulator which is not the same as the two other pairs (2F120) also has a slightly different pattern. This set of profiles encouraged us to continue with our search for common and uncommon profiles, arguing that similar symptom mixtures may imply similar genesis of problems and the converse. We also elected to review other places where motion sickness like symptoms occur and to compare

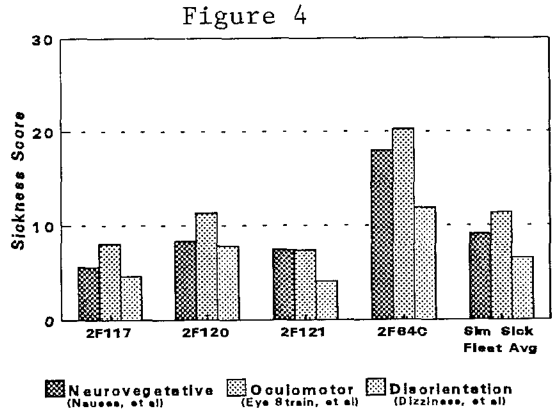
these as well.

Figure 3 shows Army helicopters and Figure 4 shows Navy helicopters. Note how most of these have similar profiles. In general the nausea component is large but the oculomotor component is the largest.

Army Helicopter Simulators
Figure 3



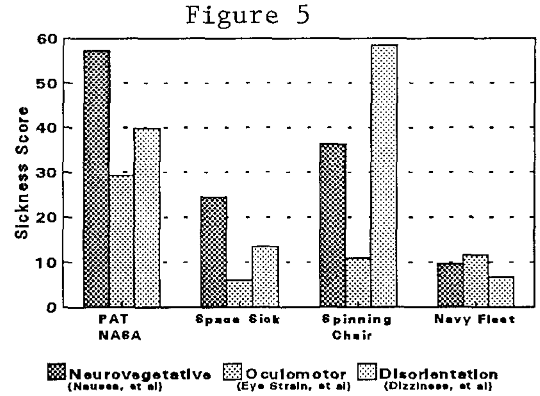
Navy Helicopter Simulators
Figure 4



When CRT based simulators are compared to dome display simulators which are also fixed base, both of these symptom incidences are reduced. We would hypothesize that the reduction in nausea is related to the moving base, and have begun studies where the moving base is turned off during operations and observed a slight lowering of this symptom complex. The relationship of eye strain to CRT displays has been commented upon by Ebenholtz (1988) and, when dome displays are used appears to be somewhat reduced. There is one system where a dome and CRT are used with the same basic flight simulator (2F120 in three locales), but these data are not yet analyzed.

Figure 5 shows the application of this methodology to space motion sickness.

Profiles of Simulator Sickness
Figure 5

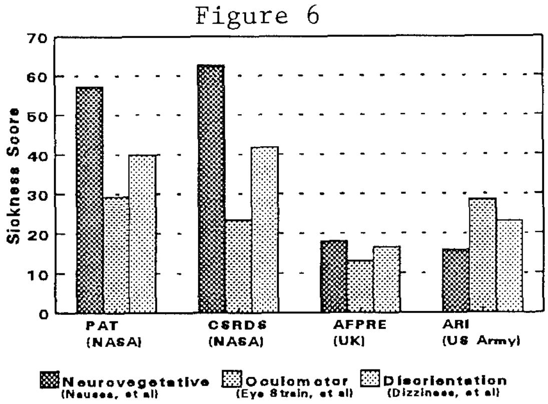


Shown in this figure is the Preflight Adaptation Trainer, a virtual reality system which is being employed to provide training in the illusory phenomena to be experienced in space in order to increase tolerance. The other data are from actual symptoms of space motion sickness reported by 85 astronauts (and other crew members) in the past several years. Also shown is a spinning chair test which is employed for pretesting for space motion sickness by NASA as well as the U.S. Navy's average for a dozen simulators (N>1500). Note first that sickness incidence is higher here than in the previous figures and we have adjusted the Y axis to a maximum sickness score of 60 versus 30. Note also that there is a very good agreement between SMS and PAT sickness and that these two are quite different from simulator sickness, but not much different from the spinning chair test. Based on these relations one might hypothesize that sickness in PAT would be more predictive of space motion sickness than either of the other two environments and that spinning chair would be better than simulator sickness.

Figure 6 shows three virtual reality systems, all of which employ head mounted displays and the NASA PAT virtual reality system.

HMD System Profiles

Figure 6

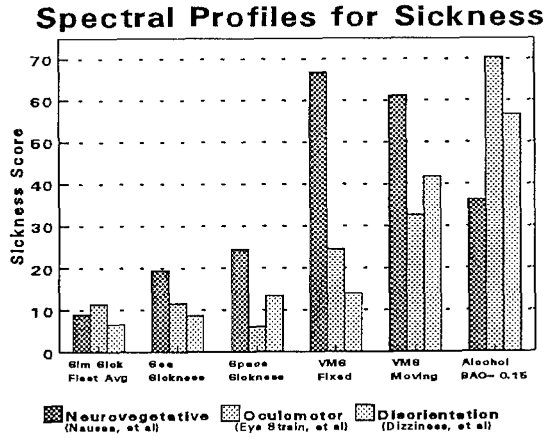


Note that the NASA Crew Station Research and Development System resembles the VR system from the UK and the NASA PAT system, but is slightly different from the U.S. Army Research Institute's VR system. Reasons for this may be uncovered by comparing the various equipment features of the different systems.

Figure 7 compares alcohol induced discomfort (at .15 mg/dL) with sea sickness and simulator sickness along with space motion sickness and sickness symptoms in the NASA VMS device.

Figure 7

Spectral Profiles for Sickness



Note that with motion on, the VMS resembles space sickness, although the magnitude of the effects are stronger in the VMS. Note that the fixed base NASA VMS resembles sea sickness. We believe that using symptom profiles such as these can shed light on the possible causes of the maladies, particularly if modifications to the devices can be made and then symptoms examined to determine whether and where changes in symptom mixture have occured.

3 DISCUSSION

Crucial to the design of VR systems is specification of the equipment parameters that will promote training effectiveness and realism, but also avoid sickness. However, the technological advances which have provided the opportunity for increased fidelity have, in turn, placed greater demands on other tolerances on simulator subsystems (e.g., responses of visual and motion base systems and their interaction). Misalignment or asynchrony among simulation modalities and channels and other failures may occasion eyestrain and other symptoms related to motion sickness. Yet, evaluators of these systems may be unaware of the complex nature of these causes and be unable to detect their presence and are, therefore, unable to provide enough information for the visual display engineer to identify and correct the problem. Needless to say, standards and specifications to address these problems are also lacking. In consequence, effective training may be compromised, and components and subsystems may be purchased that cannot be used, and so the buyer does not get good value for their acquisition dollars.

Our experience with system development suggests that assessing effects on humans usually comes late in the development process at a time when the design is virtually frozen, but VR systems are brand new and under development. We think that the first technical step in improving systems so that they do not induce sickness is to quantify, as accurately as possible, the problem(s) that are experienced by the humans who will use them. The causes cannot be determined until there is a suitable assessment of the "criterion". This

criterion in terms of which engineering characteristics will ultimately be evaluated, needs to be reliable and valid and sufficiently diverse so that differential stimulus effects can be discriminated.

In the case of simulator sickness, we have begun to measure the problem(s) experienced by the pilots as accurately as possible. We believe that this scoring system affords the opportunity to make comparisons over several different environments and stimulus conditions and provides a standardized method that can profit by widened usage in VR systems.

References

Crosby, T. N., & Kennedy, R. S. (1982). Postural disequilibrium and simulator sickness following flights in a P3-C operational flight trainer. Preprints of the 1982 Annual Scientific Meeting of the Aerospace Medical Association (pp. 147-148). Washington, DC: Aerospace Medical Association.

Ebenholtz, S. M. (1988). Sources of asthenopia in navy flight simulators. Defense Logistics Agency, Defense Technical Information Center. Alexandria, VA. Accession number: AD-A212 699.

Fowlkes, J. E., Kennedy, R. S., & Allgood, G. O. (1990). Biomedical evaluation and systems-engineering for simulators (BESS). Paper presented at the International Training Equipment Conference and Exhibition (ITEC). Birmingham, England.

Havron, H. D., & Butler, L. F. (1957). Evaluation of training effectiveness of the 2FH2 helicopter flight trainer research tool (NAVTRADEVCEN 1915-00-1). Port Washington, NY: Naval Training Device Center.

Kennedy, R. S., Berbaum, K. S., Lilienthal, M. G., Dunlap, W. P., Mulligan, W., & Funaro, J. (1987). Guidelines for alleviation of simulator sickness symptomatology (NTSC-TR-87-007). Orlando, FL: Naval Training Systems Center. DTIC # AD A182 554

Kennedy, R. S., Lane, N. E., Berbaum, K. S., & Lilienthal, M. G. (1993). Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. International Journal of Aviation Psychology, 3(3), 203-220.

Kennedy, R. S., Lilienthal, M. G., Berbaum, K. S., Baltzley, D. R., & McCauley, M. E. (1987). Symptomatology of simulator sickness in 10 U.S. Navy flight simulators (NTSC-TR-87-008). Orlando, FL: Naval Training Systems Center.

Kennedy, R. S., Tolhurst, G. C., & Graybiel, A. (1965). The effects of visual deprivation on adaptation to a rotating environment (NSAM 918). Pensacola, FL: Naval School of Aerospace Medicine.

Lane, N. E., & Kennedy, R. S. (1988). A new method for quantifying simulator sickness: Development and application of the simulator sickness questionnaire (SSQ) (Tech. Rep. EOTR 88-7). Orlando, FL: Essex Corporation.

Lentz, J. M., & Guedry, F. E. (1978). Motion sickness susceptibility: A retrospective comparison of laboratory tests. Aviation, Space, and Environmental Medicine, 49(11), 1281-1288.

McNally, W. J., & Stuart, E. A. (1942). Physiology of the labyrinth reviewed in relation to seasickness and other forms of motion sickness. War Medicine, 2, 683-771.

Miller, J. W., & Goodson, J. E. (1960). Motion sickness in a helicopter simulator. Aerospace Medicine, 31, 204-212.

Money, K. E. (1970). Motion sickness. Psychological Reviews, 50(1), 1-39.

Simulator Sickness Field Manual Mod 4 (1989, October). Naval Training Systems Center, Orlando, FL.

Virtual Reality Annual International Symposium (1993). Virtual Reality Technical Committee of the IEEE Neural Networks Council, September 18-22. Seattle, WA.

Wendt, G. R. (1968). Experiences with research on motion sickness (NASA Special Publication No. SP-187). Pensacola, FL: Fourth Symposium on the Role of Vestibular Organs in Space Exploration.

Wiker, S. F., Kennedy, R. S., McCauley, M. E., & Pepper, R. L. (1979). Reliability, validity, and application of an improved scale for assessment of motion sickness severity (USCG Tech. Rep. No. CG-D-29-79). Washington, DC: U.S. Coast Guard Office of Research and Development.

Wiker, S. F., Kennedy, R. S., McCauley, R. E., & Pepper, R. L. (1979). Susceptibility of seasickness: Influence on hull design and steaming direction. Aviation, Space, and Environmental Medicine, 50, 1046-1051.

Wiker, S. F., & Pepper, R. L. (1978). Change in crew performance, physiology and affective state due to motions aboard a small monohull vessel: A preliminary study. U.S. Department of Transportation, United States Coast Guard (Accession No. AD A077317). Washington, DC: Office of Research and Development.

Wiker, S. F., Pepper, R. L., & McCauley, M. E. (1981). A vessel class comparison of physiological, affective state, and psychomotor performance at sea. In J. C. Guignard & M. M. Harbeson (Eds.), Proceedings of the International Workshop on Research Methods in Human Motion and Vibration Studies. New Orleans: Naval Biodynamics Laboratory.

TABLE 1
Computation of SSQ Scores

Symptom (Scored 0,1,2,3)	N Nausea	O Oculomotor	D Disorientation
	Weights for Symptoms		
General Discomfort	1	1	
Fatigue		1	
Headache		1	
Eye Strain		1	
Difficulty Focusing		1	1
Increased Salivation	1		
Sweating	1		
Nausea	1		1
Difficulty Concentrating	1	1	
Fullness of Head			1
Blurred Vision		1	1
Dizzy (Eyes Open)			1
Dizzy (Eyes Closed)			1
Vertigo			1
Stomach Awareness	1		
Burping	1		
Total	[1]*	[2]	[3]

Score

N = [1] x 9.54

O = [2] x 7.58

D = [3] x 13.92

TS = [1] + [2] + [3] x 3.74

*Total is the sum obtained by adding the symptom scores. Omitted scores are zero.

On the Feasibility of Virtual Environments in Medicine

A.C.M.Dumay, Ph.D. and G.J. Jense, Ph.D.

TNO Physics and Electronics Laboratory, High Performance Computing
P.O.Box 96.864, 2509 JG The Hague, The Netherlands

(A.C.M.Dumay@fel.tno.nl G.J.Jense@fel.tno.nl)

Abstract

Virtual Environments allow a human to interact with a (computer) system in such a way that a high level of presence in a computer-synthesised world is experienced. In principle, all human senses are involved with the interaction. Many applications may benefit from this type of human-machine interfacing, however, little have emerged so far for medicine. In this paper we elaborate on some realistic potential applications of Virtual Environment technology in the field of medicine. These applications can be found in education/training, therapy, surgery, rehabilitation, diagnosis, telemedicine and biomechanics. The value to be added to these applications by VE technology lies in the fact that patient data or patient models may be moderated to the physician in a more intuitive and natural manner. Despite these potentials, the short-term feasibility of these applications can be put into question for various reasons. Firstly, the current generation of display devices have a resolution that may show to be too low to achieve a sufficiently high degree of realism for medical applications. Secondly, there are no commercially-available actuators for tactile and force feedback which the physician desperately need for the simulation of the contact with the (virtual) patient. Thirdly, the enormous computing power required for these applications needs (yet) a considerable investment. With these limitations in mind, we believe that we are at the cradle of a whole new generation of VE applications in medicine.

1. Introduction

Visualisation in medicine dates back to times long before Röntgen acquired his first image and has since then been a prime subject in, e.g., diagnostic radiology. With the advent of X-ray techniques a whole new era in medicine evolved, in which the interpretation of visualised patient data became the main topic.

In radiology five levels of information processing can be distinguished: i) image acquisition; ii) image reconstruction, iii) image processing, iv) (interactive) visualisation and v) image interpretation. At present, images can be acquired in 2-D and 3-D on the basis of three physical phenomena: i) transmission, ii) emission and iii) reflection, depending on the type of sensors used and the type of information to be collected from the patient. Visualisation of data is required at each level and needs special attention and consideration, since by visualisation the information collected from the patient is moderated to the physician. The introduction of computer techniques has made it possible to visualise and manipulate patient data for various purposes. Real-time image processing was made available as a tool to physicians and they have more or less adapted to a new way of handling and looking at images of patients. Now that High Performance Computing (HPC) technology (distributed and parallel computing) is maturing, a whole new field of visualisation applications in medicine can be explored: Virtual Environments.

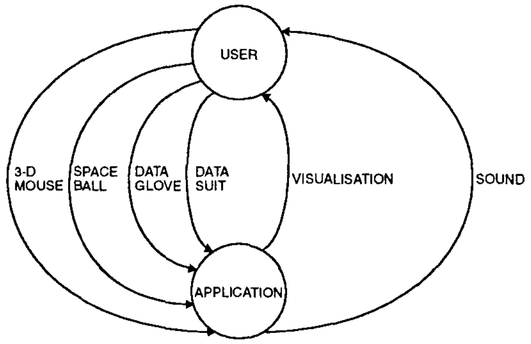


Fig.1 The human interacts with the machine through a DataSuit, DataGlove, SpaceBall or 3-D Mouse after computer stimulation of sight and hearing.

Virtual Environment (VE) is the term used by academic researchers to describe a form of human-machine interaction where the human is immersed in a world created by the machine, which is usually a computer system. Other terms in use for indicating this type of interface include cyberspace, telepresence, mirror world, artificial reality, augmented reality, wraparound compuvision and synthetic environments. In principle, all five human senses (sight, hearing, touch, taste and smell) are involved with the immersion in such a way that there is stimulation by the machine [1]. The human responds to the system by actuating peripheral sensors. The human absorbs most information by sight. Hearing comes in the second place, and touch in third. Motoric activation, speach and head/eye movements are exploited when it comes to responding to the presented information. At present, peripheral sensors are based on the DataSuit, DataGlove, SpaceBall and 3-D Mouse, while sight and hearing are the most prominent of the senses involved. Visual depth is perceived from stereo images of objects at small distances (<10m). The static phenomena related to sensing visual depth at large distances (>10m) are relative position, shading, brightness, size, perspective and texture gradient, while the single important dynamic phenomenon is motional parallax.

The term VE is derived from the term Virtual Reality (VR), which was first launched by Jaron Lanier. We also use the term VE to emphasise the embeddedness of the human in the virtual environment synthesised by the machine. In Figure 1 we illustrate the interaction between the user and the application running on the computer system.

Physicians have traditionally been very sceptical about technological innovations and it is assumed that VE developments will not lack criticism. Despite the cynicism, the word *cyber-radiology* has already been used by radiologists, suggesting that there is some measure of interest in the medical area: "What radiologist can salvage from this dark edge of the computer age are insights into the work they do and, with the help of a breakthrough or two, radical new ways of diagnosing disease" [2].

In this paper we elaborate on realistic potential applications of VE technology in medicine, including education/training, therapy, rehabilitation, surgery, diagnosis, telemedicine and biomechanics. We conclude this paper with some wishful thinking.

2. Education and training

Virtual Environment technology can already be found in training and simulation systems [3]. This is by no means surprising, since the human learns best by actively committing itself to the learning task involving as many senses as possible. The ability to interact with the virtual environment rather than just with the system makes VE training and simulation systems more appreciated than multimedia systems like interactive CD and interactive video [4]. VE training and simulation systems can give the human an artificial experience with intrinsic educational benefits. In order to provide simulation-based medical training facilities, innovative and technically demanding concepts and techniques are needed for providing natural and high quality interaction between the user and the machine. High performance computing technology is an essential ingredient of realistic visualisations of the medical field.

There are four problem areas anticipated: i) the resolution of the display devices; ii) the performance of such a system in terms of computation time; iii) the availability of medical instruments interfaced to the computer system, and iv) cost aspects. To visualise anatomical structures with an acceptable level of detail an image resolution of 1kx1k pixels or even higher may be required, although this issue is still open to discussion [5]. The rendering of a realistic anatomical model with textures, variable level of detail and simultaneously allowing for human interaction with the model requires GFLOPS of computing power. At present, there are hardly any medical instruments available that can be interfaced to computer systems. Especially tactile and force feedback to these instruments need further research and development. In literature, a first prototype of a (surgical) instrument has been reported [6].

2.1 Learning on anatomy, physiology and pathology

Learning on the (human) anatomy and physiology in VE is one educational application. Anatomical models are already available from medical text books and the physiology of the various organs is also well-documented. A mathematical description of the anatomy and physiology of body organs are to be stored in a database ready for visual rendering with a variable level of detail. More details can be revealed from view points close to the body surfaces and even details that are not visible with the naked eye can be shown. In order to gain in realism photographs of real textures of the body organs may be mapped onto the virtual organs, again with variable level of detail. This educational aid can be improved by simulating on pathologies as well and giving the user the ability to take the virtual patient apart and put it together. Optionally, the scoring of individual users to an educational programme can be recorded.

2.2 Medical emergency-room training

The emergency room of a hospital is a theatre that can only function properly when medical staff is well prepared and fully informed on procedures and protocols. This requires specialist training which may be facilitated with a VE training and simulation system. In such a system, the actual emergency room can be modelled, including beds, patient tables, drawers, curtains, surgery facilities, infusion pumps etc. The drawers may contain virtual medical aids like bandages, clamps and syringes. In principle, a virtual patient (see Section 2.1) can be exposed to any injury. In an interactive VE training session an injury can be treated following a selected protocol, giving the user the ability to cure the patient or to inflict even worse injuries. In such a training session the real atmosphere in an emergency-room can be approximated. Even a certain level of stress can be induced to the user. Again, the scoring to an educational programme can be recorded automatically.

2.3 Training of ambulance staff

A relatively simple variation to a medical emergency-room training system is one suitable for training of ambulance staff. The virtual patient does not need modification, only the virtual interior of the emergency room should be replaced by a model of the ambulance interior. A simulator of this type is especially of interest for medical services in countries where ambulance staff need more education than one for acquiring a first aid certificate.

2.4 Triage training

Another variation to the emergency-room training system is one for triage training. Triage is a protocol to assess the patient condition and to decide on medical treatment under trauma conditions with limited support from medical facilities. The assessment of the patient condition under trauma conditions requires specialist medical knowledge. This is especially true in crisis or war situations where a great number of casualties may be delivered for triage in a short period. In these situations, damage to the internal organs can usually not be rated by diagnostic X-ray screening. The treatment of selected patients aims at maintaining at least a minimal functionality of vital organs, while a limited number of time-consuming surgical interventions should be weighted against the maintenance of the mentioned functionalities within a larger number of casualties. In a VE injuries can be inflicted to virtual casualties before triage protocols and the management of wounds (see e.g. [7]) can be trained. The possibilities of nuclear, biologic or chemical (NBC) damage to the casualties make triage a real challenge, recquiring thorough education and training. To efficiently and effectively train military medical staff VE training and simulation systems can be utilised. Such systems do not replace the real interaction with the patient, but may reduce the time for training, reduce the costs of training and it gives one the ability to simulate situations that cannot be found in civil medicine. In Figure 2 the idea behind triage training in VE is illustrated. The patient is observed by the trainees through binocular display devices.

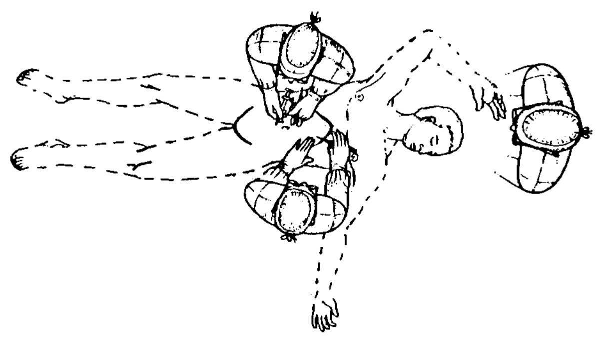


Fig.2 In a triage training session the virtual casualty is observed by the trainees through binocular display devices.

The building blocks for devising a complete VE triage training system comprise: i) a geometrical model of the human body like the MIRD-5 Adult Mathematical Model [8]; ii) an anatomical and iii) physiological model of the internal organs and muscles [9]; textural photographs of iv) conventional and v) NBC damage to the human body; a model of the mathematical physics underlying vi) rigid and vii) soft body deformations and viii) gravity; iix) triage protocols; ix) a dynamic data base; x) 3-D visualisation tools; xi) 3-D display devices and xii) interactive sensor systems.

2.5 Minimal access surgery

Minimal Access Surgery (MAS), also known as Minimally Invasive Surgery (MIS) and Laparoscopic Surgery (LS), is a new surgical procedure to operate a patient, especially in the abdomen [10]. The social and economic demand for a wide-spread use of MAS techniques leads to the requirement that large numbers of surgeons and medical students be instructed and trained. At present, a minimum of 30 hours is required for training on the procedure in live patients [11]. There are two stages of MAS training : i) video-based training and ii) training on animal models. Video-based training consists of manipulating instruments

without direct visual contact. This allows the acquisition of basic coordination and dexterity. Training on animal models is carried out under conditions that are as close as possible to these involving a human patient. A computer-based simulator which makes use of VE technology allows for an intervention being carried out on virtual human organs in a similar surgical environment as in a real operation with the possibility of introducing anatomical anomalies that one might encounter in a real operation. The user should be able to directly see the effect of manipulating the MAS instruments in the simulated image, which requires a mechanical dummy interface. In addition, facilities must be provided for trainer - trainee communications and for the overall handling of the simulation and the training system. All these mechanisms must be user-friendly and natural to allow focusing on the training task. The surgical instruments manipulated by the user should be real and their images should be synthesised in the same manner as organs. The level of realism depends mainly on the available computation power. High performance computing power permits a better discretisation of the database and so a better image, with a better behaviour of deformation models. The visualisation of simulated images can be obtained by using parallel algorithms already used in real-time synthetic image generation. The models used should be enhanced in realism by mapping 3-D textures. The organ models can be produced using solid modelling techniques from the field of mechanical engineering with a sufficient realistic appearance. The feasibility of a virtual reality surgical simulator for education and training purposes is being studies on the basis of a prototype system [12]. The drawing in Figure 3 illustrates the concept of a MAS training system.

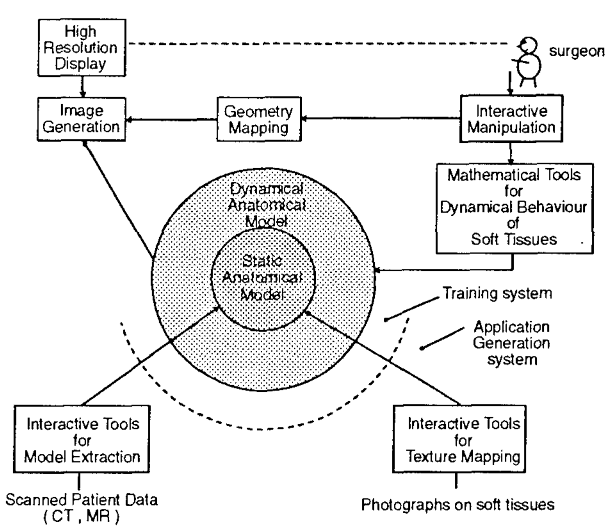


Fig. 3 The concept of a MAS training system with anatomical computer model, dynamic model of the deformation of soft tissues and computer peripherals.

3. Therapy planning

Diagnostics and therapy control are two steps in the treatment of malignant tumour cells that rely heavily on decisions made by the physician on the basis of visual impressions. One of the most important treatment methods is the radiotherapy of tumours, i.e., exposing tumours to radiation. The close vicinity to the target area of radiosensitive organs, such as the optic nerves, the spinal cord and the brain stem, often means that with conventional radiotherapy it is not possible to administer a sufficiently high dose to the tumour without inducing serious damage to the surrounding healthy tissue. With conformal precision radiotherapy planning the target area is delineated in scanned patient data and visually presented to the operator. Optimal directions of irradiation can be computed from a desired dose distribution [13]. Such a therapy planning can be carried out in a VE in which the patient can be modelled and the planning results can be shown. In Figure 4 we illustrate the localisation of a brain tumour. For this application we see similar problem areas as discussed in Chapter 2.

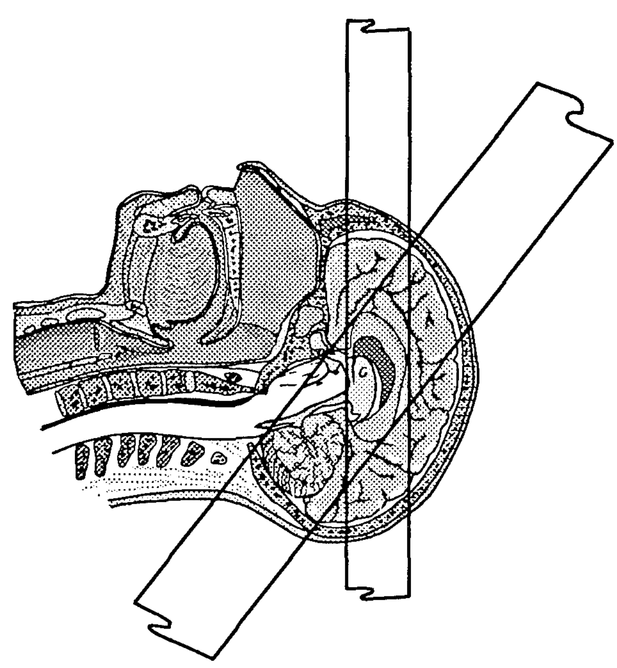


Fig.4 Optimal directions of irradiation can be computed on the basis of the location of the target area, here illustrated for the brain. The target area and the planning results can be presented in VE.

4. Rehabilitation

Another class of applications fully exploiting the presence capabilities of VE is found in rehabilitation, of which we discuss communication and therapeutic rehabilitation. For these relatively simple applications we see no limitations inflicting their feasibility.

4.1 Communication

Disabled persons lack the ability to fully participate in a society. In a graphic environment, however, there are essentially no constraints imposed to the handicapped. With a focus on the movements of hand, fingers, shoulder and face, VE technology can be applied as a means of communication with the computer system and the person's environment. A special type of communication can be found in sign language. Interactive learning of sign-language can be facilitated with a VE simulation system as suggested in Figure 5, where a sign for *Amsterdam* is demonstrated [14]. Legal signs modelled by hands attributed with data gloves can be automatically interpreted by the computer system, while incomplete signs can be shown correctly. Illegal signs should be ignored. Such a system can keep track of the score of an individual user, who might find this 'talking mirror' interesting and encouraging to use.

4.2 Therapeutic rehabilitation

VE technology can also be useful in the rehabilitation of muscles and nerve systems after surgery or accidental damage. Conventional rehabilitation sessions are often experienced as dull and VE technology may give rehabilitation a new dimension.

A patient may be motivated to therapeutic rehabilitation in interaction with a playful or even competitive VE simulation. Motoric functions can be stimulated by playing with virtual objects with a minimum of energy or effort. An additional advantage of such an approach to rehabilitation is that the patient's performance can be registered and therapy sessions be adjusted accordingly. Diseases may be recognised from response times and finger movements of a patient wearing a data glove, while dysfunctions at a perceptive level can be concluded from a patient who senses a growing conflict of sensory input. Also the relationship between the electroencephalogram (EEG) and specific cognitive activities in VE can be investigated in a therapeutic session.

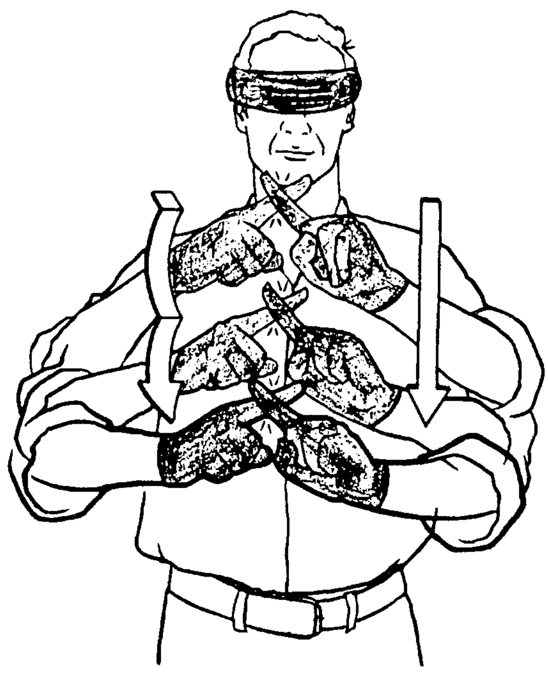


Fig.5 Interactively learning sign language in a VE 'talking mirror' may be experienced as playful and challenging. Here, a sign for 'Amsterdam' is illustrated.

5. Surgery

Besides training and simulation systems for surgery (see Section 2.5) employing VE technology, computer-assisted surgery and surgery planning may benefit from a VE user interface. The problem areas for these applications are similar to the ones discussed in Chapter 2.

5.1 Computer-assisted surgery

Computer-assisted surgery can be applied in e.g. (stereotactic) neurosurgery after careful planning of the intervention following similar 3D image processing methods as described in Chapter 3. Here, VE technology can be applied as an interface to robot-controlled surgery where an (extremely) high precision is required for micromanipulation. The surgical intervention can be carried out in VE before instructions are sent to the robot for actual intervention in the patient. The feasibility of computer-assisted surgery has already been demonstrated [15], [16].

5.2 Surgery planning

Microsurgery includes the creation of (small) blood vessels in humans. In order to study the effect of a surgical intervention animal studies are often performed prior to the actual intervention in man. To lessen the need of animal studies and to improve surgical protocols VE simulation systems can be applied to plan a surgical intervention. The need of planning and careful consideration is also present in cosmetic surgery. A graphical model can be derived from scanned patient data, and the effect of surgical interventions can be visualised. Progress and problem areas for craniofacial surgery planning are reported in [17]. The main problem encountered showed to be criticism on the limited ability for the end user to be involved with decision making.

6. Diagnostic radiology

Visualisation and interpretation are the main topics in diagnostic radiology. Traditionally, the radiologist interprets 2-D X-ray images of (parts of) the patient in order to support medical decision making. Since the introduction of computer tomography to the field of diagnostic imaging the radiologist has gained experience in interpreting 3-D X-ray, nuclear and magnetic resonance images. These 3-D images, however, are often visualised on a slice-by-slice basis rather than in 3-D, although really 3-D applications have already been introduced.

6.1 Visualisation in radiology

VE technology may provide a whole new repertoire of applications to diagnostic radiology. All the aforementioned applications should depart from scanned patient data from which computer models are to be extracted and represented in VE. The main advantage of moderating patient data to the physician through graphical models may be that this type of representation is closer to looking at real body organs than to looking at grey level images. At the same time, however, this is also a main disadvantage. Physicians have become familiar to screening grey tone images for faint flaws in structures, symmetry, grey level, etc., taking into account the nature of the physical phenomenon underlying the imaging. This brings us to the paradox of decision making. Beside the paradox on decision making, the problem areas mentioned in Chapter 2 apply here as well.

6.2 The decision paradox in medicine

Traditionally, there are five levels of information processing in radiology (see Chapter 1). At each level a certain amount of decision making is involved. This can be either intrinsic, extrinsic or both. Intrinsic decisions are inherent to applying automated computer techniques. The only extrinsic decisions involved should be made by the user, who interprets the data. Deriving a computer model of the patient from scanned data needs, with present day technology, all five levels of information processing and consequently the extrinsic decision making. At the highest level, i.e., graphical modelling in VE to support medical diagnosis, again extrinsic decisions are made for interpreting the data. Unless medical images can be processed fully automatically, the need for VE technology in diagnosis is limited despite the natural way of representation. Obviously, breakthroughs in these areas are yet to come.

7. Telemedicine

Telemedicine is the area where telecommunication meets medicine. Obviously, where telecommunication is applied some kind of human-system interface is needed, of which VE should be considered. Telemedicine can be applied in remote consultation (physician-physician and patient-physician). Remote diagnosis and surgery can be carried out by a specialist through giving assistance to a (non-) specialist in, e.g., a hazardous or inaccessible environment during actual procedures.

7.1 Teletriage

Teletriage is defined as a special type of teleconsultation, aiming at supporting the military physician in a war or crisis situation with on-line help from a remote medical expert. While the military physician is examining a casualty he can report his findings to the remote medical expert by voice. The medical expert himself, or a team of experts, can respond by projecting instructions through VE technology onto the eye of the military physician. This gives the military physician the opportunity of having his hands free and being instructed by specialists without the need of being a specialist himself.

7.2 Telediagnosis and telesurgery

Instructions from the specialist can be moderated to the consulting physician through VE. In this way, the normal and pathologic anatomy can be projected onto

the eye of the consulting physician while examining the patient. The same can be done for surgical instructions. Hazardous environments can be found in war situations while submarines and other navy vessels in full operation are considered to be inaccessible environments. Telemedicine can be particularly interesting to support peace keeping efforts of (united) nations in politically and military instable areas with a poor (medical) infrastructure.

Fig.6 The posture of a patient can be recognised using relatively simple sensor systems and presented in a VE to the patient, who can watch his posture from any position and correct for it.

8. Biomechanics

Virtual Environments can also be interesting for the field of biomechanics. Relatively simple sensor systems can be applied for human posture recognition, which in its turn can be applied to interactive posture correction. Figure 6 illustrates a "help yourself" posture correction session. The latter is of interest to disabled persons, sporting persons and physically active persons, and can be integrated in the process of designing optimal equipment for the individual of group of individuals. In this manner, optimal seats for cars, cockpits, military tanks, etc., can be developed.

9. Conclusions

In this paper we elaborate on applications of VE technology in medicine which we consider feasible. We identify the following application areas: i) education and training, including education on anatomy, physiology and pathology, medical emergency-room training, training of ambulance staff, triage training and minimal access surgery training; ii) conformal radiotherapy planning; iii) rehabilitation, including communication and therapeutic rehabilitation; iv) surgery, including computer-assisted stereotactic neurosurgery and planning of microsurgery and cosmetic surgery; v) diagnostic radiology; vi) telemedicine for consultation purposes and vii) biomechanics for interactive posture recognition and correction and the design of equipment. The most promising of these areas is education and training, since VE technology primarily allows the user to actively committing itself to a learning task with as many senses as possible.

The feasibility of the aforementioned applications can be put into question for various reasons. Firstly, the current generation of display devices has a resolution that may show to be too low for realistic medical applications. At present, some diagnostic applications are supposed to need a resolution of 4Kx4K pixels! Secondly, there are no commercially-available actuators for tactile and force feedback which the physician desperately need for the contact with the (virtual) patient. Thirdly, there is a need of GFLOPS of computing power. Although this is basically not a problem area, the financial investment needed to acquire that computing power may delay the developments. Finally, despite the realism that can be achieved within virtual environments, VR applications remain models of the real world: models cannot replace the real world ! This is true in general and for medicine in particular. We do not suggest applying VE training and simulation systems instead of "training on the spot", but see it as an additional educational tool, shortening the training period effectively and efficiently. We share Lanier's feeling in foreseeing a coming revolution on the use of computer models of humans and human organs in diagnosis and treatment [18]. However, to our understanding of medical decision making in diagnosis, quite some breakthroughs are required, possibly overruling the 'traditional' levels of information processing. There is great need of new data paradigms to support real-time medical visualisation.

With these limitations and problem areas in mind, we believe that we are at the cradle of a whole new generation of applications of Virtual Environment Technology in medicine.

References

[1] S.Aukstakalnis and D.Blatner, *Silicon Mirage*, Berkeley: Peachpit Press, 1992.

[2] P.L.Ogle, "Is the World ready for Cyber-Radiology ?," *Diagnostic Imaging International*, May 1993, p.5.

[3] G.J.Jense and F.Kuijper, "Virtual Environments for Adapted Trainers and Simulators," *Proc. Int. Training Equipment Conference*, pp.49-57, 1993.

[4] P.Vossen, "New Approaches to Medical Training. A patient simulator," submitted.

[5] H.Fujita, M.L.Giger and K.Doi, "Investigation of Basic Imaging properties in Digital Radiography: 12. Effect of Matrix Configuration on Spatial Resolution," *Med.Phys.*, vol.15, pp.384-390, 1988.

[6] D.Sims, "The Point Where Lines Converge", *IEEE Comput. Graph. and Appl.*, no. July, pp.7-8, 1993.

[7] J.Dolin, Th.Scalea, L.Mannor, S.Sclafani, et al., "The Management of Gunshot Wounds to the Face," *The Journal of Trauma*, vol.33, pp.508-515, 1992.

[8] International Commission on Radiation Units and Measurements, Phantoms and Computational Models in Therapy, Diagnosis and Protection, ICRU Report 48, 1992.

[9] T.Schiemann, M.Bomans, U.Tiede and K.H. Höhne, "Interactive 3D-Segmentation," *Proc.SPIE Visualization in Biomedical Computing*, vol. 1808, pp.376-383, 1992.

[10] P.Testas and B.Delaitre, *Chirurgie Digestive par voie Coelioscopique*, Paris: Éditions Maloine, 1991.

[11] M.D.Noar, "Endoscopy Simulation: A Brave New World?," *Endoscopy*, vol.23, pp.147-149, 1991.

[12] R.M.Satava, "Virtual Reality Surgical Simulator. The First Steps," In J.D.Mabrey, ed. *Proc. of the USASOC Surgeon's Conference*, 5-8 December 1992, pp.12-14.

[13] S. Webb, "Optimisation of Conformal Radiotherapy Dose Distributions by Simulated Annealing", *Phys. Med. Biol.*, vol.34, pp.1349-1370, 1989.

[14] L.Koenen, T.Bloem and R.Janssen, *Gebarentaal. De Taal van Doven in Nederland*, (*Sign Language. The Language of the Deaf in the Netherlands*,) Amsterdam: Nijgh & Van Ditmar, 1993.

[15] W.S.Ng, B.L.Davies, R.D.Hibberd and A.G. Timoney, "Robotic Surgery," *IEEE Eng. Med. Biol.*, pp.120-125, 1993.

[16] B.Preising, T.C.Hsia and B.Mittelstadt, "A Literature Review: Robots in Medicine", *IEEE Eng. in Med. Biol.*, no. June, pp.13-22, 1991.

[17] C.McEwan, R.Robb and I.Jackson, "Craniofacial Surgery Planning and Simulation: Current Progress and Problem Areas," in H.U.Lemke, M.L. Rhodes, C.C.Jaffe and R.Felix, eds., *Proc. Int. Symp. CAR '89*, pp.398-402, 1989.

[18] N.Hays, "Lanier 'Rants' about New VR Issues," *IEEE Comput. Graph. and Appl.*, no. May, p.85, 1993.

INTERFACES VOCALES

POUR LES SYSTEMES OPERATIONNELS

PAROLE

AIDE A LA FORMATION ET L'ENTRAINEMENT

DES CONTROLEURS DE TRAFIC AERIEN

F. Marque
Centre d'Etude de la Navigation Aérienne
9, Rue de Champagne
91200 Athis-Mons
France

T. Labarrère
Sextant Avionique
BP 91
33166 Saint Médard en Jalles Cedex
France

F. Neel
LIMSI/CNRS
BP 133
914 03 Orsay Cedex
France

I INTRODUCTION

Placé sous l'égide du Centre d'Etudes de la Navigation Aérienne (CENA) le projet PAROLE utilise les potentialités complémentaires d'industriels (STERIA INGENIERIE ET TELECOM, SEXTANT AVIONIQUE et VECSYS) et d'un organisme de recherche (LIMSI) quant à l'étude et la réalisation d'un outil d'aide à la formation et l'entraînement des contrôleurs aériens.

Basé sur l'utilisation concomitante d'une interface vocale (synthèse et reconnaissance de la parole) et d'un superviseur gérant le dialogue, le prototype est à même d'exploiter complètement le canal audio.

Aujourd'hui, la validation par les opérationnels des concepts d'IHM vocale, permet d'envisager une application opérationnelle du produit PAROLE dans les centres de formation des contrôleurs de la Navigation Aérienne.

Ce document présente l'architecture et les différents constituants de PAROLE avant d'en évaluer les applications futures possibles. Il précise de plus la méthodologie suivie, basée sur les principes d'étude du langage naturel.

II GENERALITES

Le trafic aérien subit des modifications de charges dont le public est régulièrement averti... Cette activité connaît globalement, une croissance continue depuis le début de son histoire ; en 1953,

Presented at an AGARD Meeting on 'Virtual Interfaces: Research and Applications', October 1993.

on parlait déjà de flux de trafic, et des problèmes d'encombrement de l'espace. Or, depuis cette date, l'augmentation a été d'environ 5 à 6 % par an.

Le travail des contrôleurs radar consiste à guider les avions dans l'espace aérien de façon à écouler le trafic en assurant la sécurité des vols.

Les échanges entre contrôleur et pilote se font en anglais ou dans la langue du pays survolé, si cette langue est une des langues de l'Organisation de l'Aviation Civile Internationale (OACI, dont le siège est à Montréal).

L'activité spécifique du contrôleur aérien repose sur les échanges d'informations suivants :

1) utilisant des liaisons de données : systèmes sol entre eux ;

2) utilisant un début de liaisons de données : systèmes bord/système sol ;

3) utilisant des systèmes d'entrée/sortie de données : contrôleur/système sol, ou pilote/système bord ;

4) utilisant la voix : contrôleur/pilote, ou contrôleur/contrôleur.

Formation des contrôleurs aériens

La formation des contrôleurs aériens français dure quatre ans. Elle se déroule alternativement à l'Ecole Nationale de l'Aviation Civile (ENAC), située à Toulouse, et dans les centres de contrôle dans lesquels ils sont affectés à leur sortie d'école (Centre de Contrôle Régionaux ou aérodromes contrôlés). Elle s'appuie en partie sur des cours magistraux, et en partie sur des simulations.

Celles-ci s'effectuent sur plusieurs systèmes, dont un grand système téléinformatique, qui permet de former simultanément plusieurs dizaines d'élèves au contrôle avec radar, à la fois à l'ENAC et dans les CRNA (Centre Régional de la Navigation Aérienne). Ce simulateur de formation au contrôle avec radar, baptisé "Simulateur CAUTRA", est en cours de renouvellement par un système plus moderne, baptisé ELECTRA (Ensemble Logiciel pour l'Enseignement du Contrôle du Trafic Aérien).

Par ailleurs, l'ENAC va se doter, pour la formation initiale des contrôleurs, de plusieurs exemplaires d'un simulateur simplifié de formation au contrôle avec radar.

Enfin, l'ENAC dispose de simulateurs de contrôle d'aérodrome (AERSIM).

Rôle des pilotes d'échos radar

Une séance de formation au contrôle avec radar sur simulateur implique en général la présence de trois catégories d'acteurs :

- un ou plusieurs élèves contrôleurs,
- un ou plusieurs instructeurs,
- un ou plusieurs opérateurs appelés "pilotes d'échos radar", ou "pseudo-pilotes".

Le rôle du pseudo-pilote est de simuler les interventions en phonie des pilotes d'avions présents dans le secteur de contrôle. Pour ce faire, il dialogue avec l'élève et commande, à l'aide d'un écran tactile, d'un clavier ou d'une souris, le mouvement des plots sur l'écran radar simulé, en fonction des instructions données par l'élève dans le cadre d'exercices enregistrés à l'avance dans le simulateur.

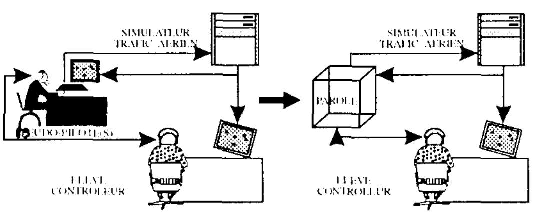
Les échanges entre l'élève-contrôleur et le pseudo-pilote se font verbalement au moyen de microphones et d'écouteurs, en français ou en anglais suivant la nationalité des compagnies aériennes des avions simulés dans l'exercice. La situation et les mouvements des avions dans le secteur aérien peuvent être suivis par l'instructeur, le pseudo-pilote et l'élève chacun sur un écran graphique où sont simulés les échos radars.

Lors de certains exercices simples, l'instructeur peut jouer lui-même le rôle du pseudo-pilote. A l'inverse, des exercices ou des expérimentations complexes, impliquant plusieurs secteurs de contrôle et plusieurs équipes de contrôleurs, peuvent nécessiter de faire appel simultanément à plusieurs pseudo-pilotes, chacun d'eux pouvant se charger de plusieurs avions simulés (jusqu'à 20 ou 30 par pseudo-pilote).

Le pilotage des échos radar représente une contrainte importante pour la formation et l'entraînement des contrôleurs, car il nécessite l'emploi de personnels bilingues et connaissant bien les procédures du contrôle du trafic aérien, pour des tâches souvent fastidieuses.

Le but de l'IHM_VOCALE PAROLE est de remplacer peu à peu les pseudo-pilotes humains dans les simulateurs de formation.

La figure suivante représente cette évolution :



III DESCRIPTION DE L'APPLICATION ET PRESENTATION DU PRODUIT

3.1 Architecture de PAROLE et organisation du projet

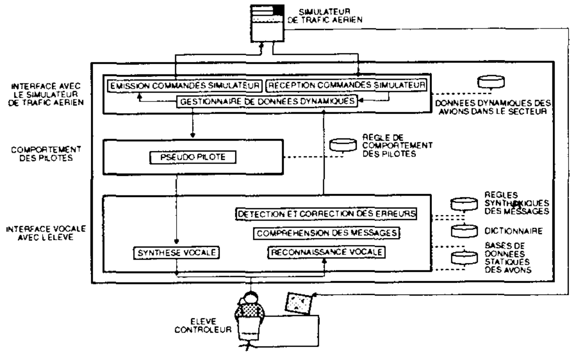
Cette substitution impose de doter PAROLE de capacités de traitement audio (Reconnaissance de la parole et synthèse vocale) mais aussi de gestion de dialogue permettant de simuler les comportements "pilote".

La figure suivante présente l'architecture globale du produit et souligne les différentes interfaces :

- interface vocale avec l'élève

- simulation du comportement pilote

- interface avec le simulateur de trafic aérien.



Les différents intervenants sont :

1) le LIMSI (Laboratoire d'Informatique pour la Mécanique et les Sciences de l'Ingénieur) a développé, en collaboration avec le CENA, la première maquette de laboratoire, puis a fourni son expertise dans le domaine gestion des dialogues, pour l'utilisation en simulateur et la gestion des parties reconnaissance de la parole/synthèse vocale,

2) STERIA a développé la partie Poste pilote et, avec la participation d'un ingénieur du CENA, la partie terminal simulateur permettant d'interfacer PAROLE avec le simulateur,

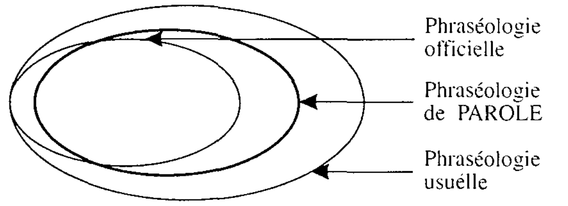
3) SEXTANT a développé la partie Terminal Vocal. Pour ce faire VECSYS a réalisé une carte de reconnaissance vocale bilangue temps réel, basée sur le système DATAVOX.

Le CENA gérant le projet et apportant son expertise tant dans le domaine opérationnel que dans celui de l'ergonomie du Contrôle Aérien.

3.2 Phraséologie

Dans le contexte opérationnel, les contrôleurs et les pilotes utilisent deux langues : le français et l'anglais. Cette spécificité impose au produit PAROLE de reconnaître la langue utilisée et de synthétiser la réponse dans cette même langue, ou de "décider" de la langue à utiliser en fonction de la compagnie aérienne.

La phraséologie officielle aéronautique, très structurée et précise n'est cependant pas utilisée de manière très pure dans le monde réel du contrôle aérien. Celle définie dans le cadre du projet PAROLE est plus riche que la phraséologie officielle, permettant de s'affranchir un peu lors des simulations du caractère rigoureux de cette dernière, tout en imposant aux élèves contrôleurs un respect minimum quant à la syntaxe à utiliser. La figure suivante illustre cette caractéristique.



3.3 Partie "Terminal vocal"

3.3.1 Fonctions du Terminal vocal

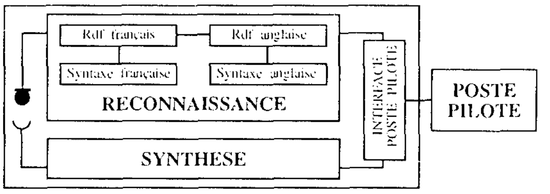
Les principales fonctions assurées par le terminal vocal (TV) sont :

1) reconnaissance des commandes prononcées par le locuteur et de la langue utilisée,

2) synthèse vocale dans la langue utilisée par le locuteur,

3) dialogue avec le système Poste Pilote (PP).

La figure suivante précise l'architecture de la partie "Terminal Vocal".



3.3.2 Modes de fonctionnement du Terminal vocal

Le Terminal vocal peut fonctionner en mode autonome ou connecté au Poste Pilote.

a) Mode autonome

Ce mode permet essentiellement d'effectuer une évaluation de la reconnaissance de la parole.

Le système propose la création de fichiers de statistiques (sur un locuteur ou moyennés sur plusieurs locuteurs) qui permet de relever les erreurs de reconnaissance systématiques et de notifier les performances de reconnaissance.

b) Mode connecté

Ce mode permet le fonctionnement avec le Poste Pilote dans différents sous-modes avec des options de statistique ou d'enregistrement garantissant une souplesse d'utilisation.

3.3.3 Phase d'apprentissage

De base le système de reconnaissance vocale est monolocuteur. Ainsi avant d'utiliser le TV, le locuteur doit effectuer une phase d'apprentissage supervisé. Pour chaque langue, celui-ci prononce les mots du vocabulaire ainsi que des phrases cohérentes qui permettent d'acquérir les références acoustiques de sa voix.

3.3.4 Caractéristiques du Terminal vocal

Les principales performances et caractéristiques techniques du terminal vocal sont les suivantes :

- reconnaissance vocale bilingue (Français/Anglais) simultanée avec alternat,

- vocabulaire syntaxé de 2 x 280 mots incluants des digits,

- commandes de 4 à 25 mots enchaînés,

- taux de reconnaissance global au niveau de commande supérieur à 95 %. Ce qui correspond en pratique aux taux observés dans le dialogue réel entre contrôleurs et pilotes d'avion,

- temps de réponse de reconnaissance inférieur à 300 millisecondes.

- synthèse vocale en français ou en anglais, fonction de la compagnie aérienne et de la langue utilisée dans la question posée. Possibilité de générer 9 voix différentes augmentant ainsi le réalisme de la simulation.

3.4 Partie "Poste Pilote"

3.4.1 Caractéristiques du dialogue Contrôleur / Pilote

a) Structure du dialogue

Dès qu'un avion entre dans le secteur géré par le contrôleur, il signale son identité (indicatif) à ce dernier, déclenchant ainsi un dialogue qui va durer pendant toute la traversée du secteur. Le contrôleur assure un dialogue avec chaque avion présent dans son secteur; ce dialogue peut être intense ou se limiter à l'initialisation et à la terminaison; en général pour chaque avion le dialogue est composé de plusieurs communications et chaque communication est elle-même composée de plusieurs échanges, un échange étant un message du contrôleur suivi d'un message du pilote ou inversement.

1) établissement
Contrôleur <----(1)---- Pilote
----(2)---->
(1) : le pilote signale son indicatif
(2) : accusé de réception + [instruction ou question]

2) échange
Contrôleur <----(3)---- Pilote
----(4)---->
(3) : question, ou instruction
(4) : réponse, ou collationnement
Contrôleur <----(5)---- Pilote
----(6)---->
(5) : initiative du pilote
(6) : réponse

3) fin
Contrôleur ----(7)----> Pilote
(7) : "au revoir", ou "contactez..."

b) Structure du message

Les messages peuvent être simples ou composés. Un message simple correspond à un seul concept sémantico-pragmatique de la tâche et peut être une instruction, une question, une information ou un message de gestion du dialogue (répétez,...)

types de message:

* instruction
* question
* information

catégories de message :

* cap, absolu ou relatif,
* niveau,
* vitesse, en noeuds ou en Mach,
* balise,
* taux d'évolution,
* contact secteur suivant,
* gestion de dialogue, répétition, confirmation.

Une catégorie de message se compose de plusieurs constituants : l'action, le sujet, les paramètres, le mode d'exécution, la butée (limite d'exécution), le délai d'exécution et éventuellement un commentaire. Les valeurs que peuvent prendre les différents constituants du message changent selon la catégorie du message.

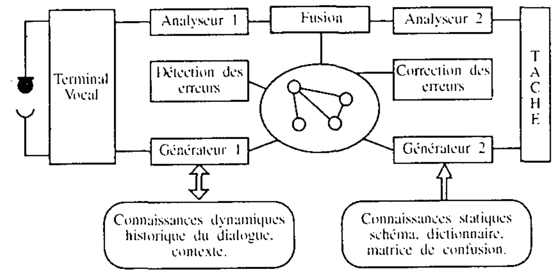
c) Lexique

Le lexique se compose d'un sous-lexique stable, et d'un sous-lexique non stable. Les mots du sous-lexique stable ne dépendent pas d'un exercice particulier, ce sont les mots clefs (maintenez,..., cap,... etc,) et les paramètres tels que les chiffres et les lettres. Les mots du sous-lexique non stable dépendent de l'exercice courant, ce sont les noms propres tels que les noms de compagnies, les noms de balises et les noms de stations.

Tout ce qui vient d'être décrit concerne le langage et le dialogue pour les deux langues : français et anglais.

3.4.2 Système de gestion du dialogue PSP

La figure suivante présente l'architecture de la partie Poste Pilote.



Le module du dialogue Poste Pilote (PP) est le module central du système : il a la charge de coordonner le fonctionnement des différents modules et principalement l'interprétation contextuelle des messages provenant du locuteur... Au cours du dialogue, le système acquiert des informations qu'il intègre au fur et à mesure dans son réseau (en créant ou éliminant, en corrigeant ou modifiant d'autres éléments). Le système dispose d'une structure sous forme de réseau de schémas, appelé réseau du dialogue, qui représente l'état courant du dialogue : on y trouve toutes les informations acquises par le système depuis le début de la négociation d'une requête ; dès qu'une action est prête, c'est-à-dire dispose de tous les paramètres nécessaires à son exécution, elle est envoyée à son destinataire (simulateur ou locuteur).

Le module traite tous les schémas indépendamment de leur provenance.

Le réseau de schémas fourni par l'analyseur ("Analyseur1" pour les messages du locuteur et "Analyseur2" pour les messages de la tâche) est ensuite fusionné dans le réseau du dialogue. Ensuite les différents modules de traitement sont déclenchés dans l'ordre suivant :

1) Détection et correction d'erreurs.

2) Génération de messages pour le simulateur et pour le locuteur.

3) Mise à jour du réseau du dialogue et de la base des connaissances.

Chacun de ces modules accède au réseau du dialogue pour récupérer des informations et en déposer d'autres. A l'initialisation du système, le réseau du dialogue est vide ; il est initialisé par le premier message provenant du locuteur ou de la tâche.

3.4.3 Représentation des connaissances

Le modèle proposé est fondé sur la théorie des schémas. Ce concept a été introduit par Minsky (Min 75) dans le cadre de la vision et repris ensuite dans de nombreux travaux avec des interprétations

très différentes (Bob 77). La philosophie des schémas est de représenter chaque objet ou concept par sa description (ensemble de champs) : puis lors de la phase de compréhension, de vérifier dans quelle mesure un texte ou un message peut se rapporter à la description ainsi formalisée.

Les connaissances sur le langage, les actions, l'association entre messages et actions réalisé par la tâche, les contraintes régissant les composants des messages et le contexte, les messages du système destinés à l'usager et associé à différents contextes et le contexte de la tâche.

On définit une catégorie comme un sous-ensemble d'actions, à chaque catégorie correspond donc un sous ensemble de messages. A chaque catégorie de messages, on associe un schéma. Toutes les connaissances nécessaires à l'analyse et à la compréhension du message sont représentées dans le schéma. Pour la définition et la manipulation de schémas on a défini un noyau de langage de schémas. Dans un schéma, on distingue plusieurs types de champs : des champs instanciés directement par analyse du message, des champs constants dont la valeur est héritée directement du schéma descripteur, des champs instanciés à partir du contexte et des champs instanciés à l'aide d'une fonction qui fournit une valeur calculée à partir des valeurs des autres champs. Il y a aussi des champs associés à des sous-schémas permettant d'établir des liens avec d'autres schémas. A chaque champ, sont donc associés des indicateurs servant de directives pour l'instanciation. Le schéma renferme aussi la définition (déclarative) des règles de contraintes de validité, les commandes simulateur (buts) ainsi que les messages de retour vers le locuteur.

Le système traite en simultané des messages en anglais et en français. A chaque catégorie représentant un ensemble de messages en français correspond une catégorie représentant l'équivalent en anglais. Il y a donc autant de schémas pour le français que pour l'anglais.

A chaque avion présent dans le secteur aérien du contrôleur, on associe toutes les informations le concernant. Ceci permet au système d'avoir une image de l'état de l'univers de chaque avion. C'est ce qu'on appelle le *contexte de la tâche*. Le système garde une trace des messages traités (historique du dialogue qui est différent du réseau du dialogue).

3.4.4 Analyse des messages

Dans PAROLE, l'analyse d'un message provenant du locuteur est dirigée directement par la tâche, c'est-à-dire que le message est analysé en sachant qu'il doit correspondre à une action ou à un concept prédéterminé et connu du système ; on utilise des heuristiques permettant d'associer le message prononcé à l'un des concepts (décrit par les schémas) de la base de connaissances. L'analyse d'un message consiste alors à déterminer sa catégorie puis à instancier le schéma correspondant à cette catégorie. L'instanciation consiste à collecter à partir du message des informations pour les inclure dans le schéma. L'exploration du message est orientée par les directives correspondant à chaque champ. La langue du message est donnée par le système de reconnaissance, ce qui permet d'activer les schémas et le dictionnaire de la langue concernée.

L'analyseur de messages reçoit en entrée une suite de mots et délivre en sortie une structure représentant le message sous forme d'un réseau de schémas.

3.4.5 Détection des erreurs

La reconnaissance automatique de la parole introduit dans la compréhension des messages un paramètre perturbateur dû au non-déterminisme de la reconnaissance. Le système doit être capable de détecter les erreurs commises par la partie reconnaissance de la parole ou éventuellement le locuteur, et extraire un sens d'un message même incomplet (partiellement reconnu) : il doit également, le cas échéant, être susceptible d'en détecter l'incohérence ou l'ambiguïté, et y remédier afin de minimiser le nombre d'échanges et, ainsi, éviter de rejeter le message dans son intégralité, chaque fois que cela est possible (Ber 84).

Deux principes sont utilisés pour le contrôle de cohérence d'un message : la limitation du domaine de variabilité et la redondance de l'information (How 89) : par exemple, dans le message "descendez au niveau 230", l'information apportée par "descendez" est inclue dans l'information "niveau 230" et dans la connaissance du niveau actuel. Trois types de contraintes ont pu être distingués :

- Contraintes globales pour toutes les catégories,

- Contraintes régissant les champs d'une même catégorie,

- Contraintes propres à chaque champ du schéma.

Ces contraintes sont intégrées de façon déclaratives dans le schéma lors de sa définition. Le module de contrôle de validité parcourt tous les schémas du

réseau pour vérifier la validité de chacun de ces schémas, en déclenchant les règles de contraintes sémantico-pragmatiques associées à chaque schéma. Si une règle n'est pas vérifiée, le champ concerné est marqué en conséquence.

3.4.6 Correction des erreurs

Des tests effectués avec quatre locuteurs sur un corpus de 50 phrases prononcées 2 fois, ont montré que l'erreur la plus fréquente du système de reconnaissance (plus de la moitié des cas : 56 %) était une erreur de confusion. Les tests ont ainsi permis de définir une matrice de confusion entre les mots comportant pour chaque mot une liste, ordonnée selon la fréquence de confusion, de tous les mots susceptibles d'être confondus avec ce mot. La matrice n'est pas symétrique.

Exemple :

(9 : 2 noeuds niveau)
(7 : 5)
(cap : 4)

Le système cherche dans le schéma erroné un champ bien instancié (îlot de confiance) voisin (prédécesseur ou successeur) d'un champ erroné ; à partir du mot correspondant à ce champ dans le message, on essaie de retrouver le mot correct en utilisant, d'une part, la syntaxe locale et, d'autre part, la matrice de confusion. Le mot non compatible est remplacé par un mot avec lequel il peut être confondu et qui figure dans la liste autorisée des prédécesseurs ou successeurs du mot considéré comme îlot de confiance, et l'instanciation est recommencée sur tout le message ce qui permet de prendre en compte la répercussion due à cette correction.

Message prononcé : "tournez gauche cap 2 3 0"
Message reconnu : "tournez gauche 4 2 3 0"
Après la correction du sujet, le message reconnu devient "tournez gauche cap 2 3 0".

3.4.7 Génération de messages

Les messages à générer sont de 4 types :

- réponse à une question posée par le locuteur,

- messages de confirmation à la suite d'une action demandée par le locuteur et réalisée par le système,

- messages de gestion du dialogue (répétition, rappel, attente, accusé de réception),

- questions qui sont en général des demandes de précision sur le message précédent du locuteur, permettant la compréhension du message ponctuel.

Le système parcourt tout le réseau et empile dans une liste tous les messages à envoyer. Les messages à générer sont déterminés par l'état du schéma : s'il est erroné le système génère une question, ainsi toutes les informations manquantes vont être demandées par le système auprès du locuteur ; l'état des schémas concernés est mis en attente d'intervention du locuteur. S'il s'agit d'un schéma satisfait, le système génère un message dont la nature est précisée dans le schéma descriptif : il peut s'agir d'une réponse à une question, d'un accusé de réception, d'une répétition ou d'un collationnement.

IV ERGONOMIE DU PRODUIT

4.1 Synthèse vocale

Lors d'un entraînement des contrôleurs aériens, l'environnement de simulation doit être le plus proche possible de la réalité. Pour ce faire, le système de synthèse vocale a été déterminé quant à sa capacité de générer des voix différentes présentant plusieurs intonations.

De plus, on part du principe que, plus une simulation est complexe, plus le langage utilisé est riche, en vocabulaire comme en tournures de phrases. PAROLE offre la possibilité de maîtriser la complexité du langage, permettant ainsi aux instructeurs de régler cette complexité sur la complexité des simulations. C'est une partie très importante pour l'apprentissage des tâches du contrôleur aérien.

4.2 Reconnaissance de la parole

Les conditions d'enregistrements sont déterminantes dans les mesures de taux de reconnaissance. Pour l'application PAROLE, le contexte peut être jugé "favorable" compte tenu :

- de l'habitude que les contrôleurs aériens expérimentés ont à parler dans un microphone,
- de l'utilisation d'un alternat microphone au début et à la fin des messages contrôleurs,
- de l'ambiance peu bruitée des salles de formation, non représentative de celles de contrôle.

V PERSPECTIVES D'APPLICATION DE PAROLE

De base, PAROLE a été défini et réalisé pour se substituer aux pilotes dans la formation et l'entraînement des contrôleurs aériens. Les différentes évaluations opérationnelles ont permis, avec la participation de contrôleurs de Roissy-Charles de Gaulle, de valider les principes ergonomiques du produit et de vérifier les performances globales de reconnaissance.

De manière à valider plus en profondeur le produit, PAROLE sera installé en 1994 au CRNA de Bordeaux.

La solution développée dans le cadre du projet PAROLE permet d'augmenter :

- la performance de l'intervention de l'instructeur, en assurant une partie des exercices de façon autonome,

- la qualité de la formation : le système permet au contrôleur aérien de s'exercer autant de fois qu'il le veut sans trop dépendre de son instructeur.

Moyennant quelques modifications, ainsi que des mises à jour, ce produit peut être adapté :

- à d'autres simulateurs, que ce soit pour le contrôle aérien ou pour le pilotage des avions,

- à d'autres vocabulaires ou syntaxes,

- aux autres langues de l'OACI.

Il peut également être utilisé, du fait du traitement multilingue et de la compréhension des langages de type opératif, dans d'autres secteurs d'activité, tels que transport ferré (formation des conducteurs, des aiguilleurs...), transport maritime (formation des pilotes, des contrôleurs maritimes...).

Ces applications peuvent concerner aussi bien les domaines civils que militaires.

Dans le domaine de la navigation aérienne, PAROLE sera utilisé par le nouveau simulateur de trafic aérien à l'ENAC, pour la formation des contrôleurs.

L'avenir de la Navigation Aérienne commence à se dessiner; il est fait d'échanges de données entre le système sol et le système bord, permettant une plus grande richesse d'informations, et une plus grande fluidité du trafic.

De ce fait, des moyens de dialogue entre le contrôleur et le système sol sont à l'étude, ce système envoyant ensuite au système bord les instructions de contrôle.

Du fait que tous les avions ne seront pas équipés immédiatement de ces nouveaux systèmes, le contrôleur aérien devra toujours donner également son instruction de contrôle pour le canal vocal VHF.

Une extension possible de PAROLE sera par exemple de le coupler à un calculateur interrogeant l'avion sur son équipement, et choisissant ainsi si le message doit être envoyé au système bord directement, avec retour devant les yeux du pilote, par sécurité, ou par le canal audio VHF, vers les haut-parleurs du poste de pilotage.

Bibliographie

C. Bailleul (86) : "*Evaluation d'un système de reconnaissance vocale dans les tâches de contrôle aérien*", Actes des 16èmes Journées d'Etude sur la Parole.

(BER 84) D. Béroule, F. Néel, "*Une approche des problèmes liés à la communication parlée homme-machine*", 4ème Congré AFCET-INRIA "*Reconnaissance des formes et intelligence artificielle*", Paris 25-26 Janvier 1984.

(Bob 77) D.G. Bobrow et al., "*GUS : a frame-driven dialog system*". Artificial Intelligence 8(2),155-173, 1977.

P. Falzon (82) : "*Les communications verbales en situation de travail - Analyse des restrictions du langage naturel*", Rapport technique INRIA nx 19.

(Fal 83) P. Falzon "*Understanding a technical language, A schema-based approach*". Rapport technique NO 237 INRIA 1983.

(Gau 82) J. -L. Gauvain, "*Reconnaissance de mots enchaînés et détection de mots dans la parole continue*". Thèse de troisième cycle, Paris XI, Juin 1982.

(Ger 93) M. GERLACH, R. ONKEN, "*Speech Input/Output as interfaces devices for communication between aircraft pilots and the pilot assistant systeme "CASSY"*, joint ESCA-NATO/RSG ETRW on *"Applications of speech technology"* Lautrach, 16-17 Sept 93.

(How 89) J. Howes "*Dialogue supervision and error correction strategy for a spoken man/machine interface*" The structure of Multimodal Dialogue. ed M. Taylor, F. Néel, D. Bouwhuis. North.Holland 1989.

(Lev 80) S. E. Levinson and K.L. Shipley, "*A conversational-mode airline information and reservation system using speech input and output*". The bell systems Technical Journal 59 (1980), 119-137.

F. MARQUE, S. TRINH, D. ABADIE, *"Exemple d'application à la formation des contrôleurs de trafic aérien : projet PAROLE"*, Ecole d'automne 1992.

F. MARQUE, *"PAROLE : du prototype à l'industrialisation"* Ergo IA 1992.

(Mat 92) K. MATROUF, F. NEEL *"Use of upper level knowledge to imporve human machine interaction"*, 2nd Venaco workshop *"the structure of multimodal dialogue"* ESCA ETRW. Acquafredda di Maraica, Italy, September 16-20-1991.

(Mat 90) K. Matrouf, "*Un système de dialogue orienté par la tâche*", Thèse de l'université Paris XI, Septembre 1990.

(Min 75) M. Minsky, "*A framework for representing knowledge*" in P.H. WINSTON, éditeur THE PSYCHOLOGY OF COMPUTER VISION, MC GRAW HILL, 1975.

(HWi 77) J. Wolf and W. Woods "*The Hwim speech understanding system*", in Trends in speech recognition, Prentice-Hall, Englewood Cliffs, pp 316-339, 1977.

LE GRAND ECRAN INTERACTIF: UN OUTIL DE DIALOGUE MULTIMODAL POUR LES FUTURES CABINES DE PILOTAGE

B. Barbier, E. Filiatre, I. Irigaray
SEXTANT Avionique
Rue Toussaint Catros, B.P.91
33166 Saint-Médard-en-Jalles Cedex
France

SUMMARY
The experimental moke-up described here is constituted of a large size projection screen displaying an image on which an operator acts in real time, under control of a specific dialogue software, using several control devices (speech recognizer, numeric data glove, oculometer). Various human communication channels are then simultaneousely used: vision and audition for the system-to-man flow, voice, gesture and gaze, for the man-to-system flow. Various ways of using and associating these communication channels allow to elaborate a multimodal dialogue.

SOMMAIRE
La maquette expérimentale décrite ici est constituée d'un écran de grande taille présentant une image sur laquelle un opérateur agit en temps réel, sous le contrôle d'un logiciel de dialogue spécialisé, au moyen de différents dispositifs de commande (analyseur vocal, gant numérique, oculomètre). Plusieurs canaux de communication humaine sont ainsi exploités simultanément: visuel et auditif pour le flux système-homme, vocal, gestuel et oculomoteur pour le flux homme-système. Les divers modes d'utilisation de chacun d'eux et les différentes façons de les associer permettent d'établir un dialogue à modalités multiples.

INTRODUCTION
Les performances de l'association aéronef-pilote tiennent pour une bonne part à la qualité du dialogue échangé entre le pilote et le système-avion. La puissance croissante des systèmes futurs ne pourra réellement être mise à profit qu'à condition de disposer d'une interface capable de véhiculer des messages de plus en plus riches et de plus en plus denses, sans augmenter pour autant la charge de travail globale.
Le concept du Grand Ecran Interactif qui va être décrit constitue l'une des solutions possible pour atteindre ce but. La maquette expérimentale d'interface qui a été développée a pour objectif d'évaluer les possibilités d'un tel système et constitue également un support d'étude pour l'optimisation du dialogue homme-système.

LES LIMITATIONS DES CABINES DE PILOTAGE ACTUELLES
Le travail de pilotage des aéronefs modernes ne consiste plus seulement à commander et contrôler directement le vol, mais aussi à dialoguer avec le système-avion.

L'interface homme-système des cabines de pilotage actuelles est structurée de la façon suivante:
- du système vers le pilote, les informations sont visuelles, elles sont données en planche de bord sur des instruments électro-mécaniques et sur des visualisations électroniques présentant des symbologies synthétiques ou des figurations issues de capteurs d'images .
- du pilote vers le système-avion, les commandes sont effectuées à l'aide de dispositifs manuels.

Cette structure d'interface souffre de plusieurs limitations:
1°) Seuls les canaux visuel et manuel sont mis à contribution.
2°) Le fractionnement de la planche de bord en plusieurs équipements de visualisation limite le mode de présentation de l'information visuelle: partition en plusieurs figurations disjointes, de position et de taille fixe.
3°) Les dispositifs de commande sont simples, avec peu de degrés de liberté et ne proposent qu'une unique modalité d'utilisation (par exemple bouton poussoir ou rotatif). Les commandes complexes s'effectuent alors en activant plusieurs dispositifs suivant une procédure spécifique; celle-ci est d'autant plus lourde à mettre en oeuvre que ces dispositifs sont différents et dispersées dans la cabine.
4°) Certains dispositifs de commande manuelle ou de contrôle visuel, dédiés à une fonction spécifique qui n'est que rarement activée, restent présents en permanence dans le cockpit, réduisant d'autant le volume disponible.

UNE VOIE POUR LE FUTUR: LE GRAND ECRAN INTERACTIF
L'étude des possibilités d'évolution des cabines du futur nous a amené à concevoir la maquette de Grand Ecran Interactif. Celle-ci est composée
- d'un écran unique occupant la totalité de la planche de bord,
- de nouveaux média d'entrée (système de mesure de la direction du regard, système de reconnaissance vocale, système de reconnaissance gestuelle)
- d'un synthétiseur vocal.

Cette maquette d'interface offre les avantages suivants:
- elle permet d'exploiter simultanément plusieurs canaux de communication, aussi bien comme moyen d'acquisition que comme effecteur,
- elle enrichit considérablement le contenu et la densité des messages échangés grâce aux capacités des nouveaux médias de communication qu'elle offre,
- elle permet la combinaison de plusieurs médias d'entrée selon différentes modalités pour réaliser une même commande offrant ainsi la capacité de "commandes multimodales",
- elle procure une grande diversité de présentations graphiques de l'état du système (figurations mobiles, ajustables en taille, superposables),
- elle fournit sur le même écran un retour visuel permanent de la commande en cours, offrant ainsi la capacité de "commande sur image",

Presented at an AGARD Meeting on 'Virtual Interfaces: Research and Applications', October 1993.

- elle diminue le nombre de dispositifs de commande.

CONSTITUTION MATERIELLE DU GRAND ECRAN INTERACTIF

Donnée en figure 1, cette maquette expérimentale est constituée des trois sous-ensembles suivants: les médias d'entrée ou de sortie, le processeur de gestion, la source d'images.

Médias de sortie

- Un rétroprojecteur LCD fournit des images couleur de 440 x 480 pixels sur un grand écran de 520 x 400 mm^2 occupant toute la planche de bord.
- Un synthétiseur vocal (Datavox).

Médias d'entrée

Ils sont constitués à partir des dispositifs suivants.
- Un oculomètre (NAC EMR-V) mesure la direction du regard par rapport à la tête à l'aide d'une micro-caméra analysant le reflet cornéen d'une diode infra-rouge éclairant l'oeil droit.
- Un système de reconnaissance de la parole en continu (Datavox), déclenché par détection d'activité, effectue une analyse phonétique et syntaxique du signal après l'avoir séparé en messages et en mots.
- Un dispositif à fibres optiques, équipant un gant porté par la main, permet de mesurer l'angle de flexion des deux premières articulations de chaque doigt (Data Glove de VPL Research).
- Des capteurs électromagnétiques (Polhemus 3 Space Isotrak) couplés à des émetteurs fixes donnent position et orientation de la main et de la tête.

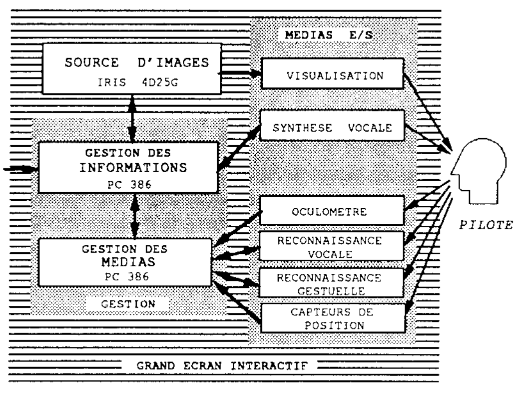


Figure 1: Interface pour dialogue multimodal.

Processeurs de gestion

- La gestion des médias est assurée par un PC 386/20 MHz; il canalise les données brutes fournies séparément par chaque média, et délivre un message multimédia au processeur de gestion du dialogue.
- La gestion intelligente du dialogue est effectuée par un autre PC 386/25 MHz; celui-ci commande la source d'images et le synthétiseur vocal.

Source d'images

La station de travail IRIS 4D25G fournit, au rétroprojecteur, les images de type TV. Renouvelées à un rythme dépendant de leur complexité (environ 8 Hz), elles sont constituées pour l'essentiel:
- en zone centrale, de fenêtres variables en taille et en position, chacune contient une figuration avionique de type déterminé, elles peuvent se superposer partiellement ou totalement suivant plusieurs plans.
- en partie inférieure, d'étiquettes représentant l'état des fenêtres (présence à l'écran) et celui des médias (marche, arrêt, panne), ainsi qu'une zone de sécurité présentant les fenêtres prioritaires, en médaillon.

MEDIAS D'ENTREE ET CANAUX DE COMMUNICATION

Regard

Sa direction est calculée en permanence d'une part à partir de l'orientation de l'oeil par rapport à la tête et d'autre part à partir de la position et de l'orientation de la tête par rapport à l'écran. La direction est indiquée par un symbole spécifique, mobile sur l'écran; l'activité de la main a priorité sur ce canal.

Voix

Le vocabulaire est volontairement restreint à 36 mots, regroupés en messages de 1 à 3 mots.

Main

Un vocabulaire postural simple a été défini; il contient les 4 postures manuelles suivantes:
-"*désigne*": index tendu, autres doigts repliés.
-"*o.k.*": pouce levé.
-"*pris*": main fermée, doigts tendus.
-"*stop*": main ouverte, pouce replié.
Le suivi de désignation de l'index est assuré par un symbole spécifique mobile sur l'écran

Les 5 gestes suivants définissent le vocabulaire gestuel:
-"*prendre*": main proche de l'écran, puis se refermant en posture "pris".
-"*lâcher*": quitter la posture "pris".
-"*jeter*": lâcher après s'être éloigné de l'écran.
-"*rotation droite*": tourner la main de 30° sur elle-même, en posture "pris".
-"*rotation gauche*": tourner la main dans le sens inverse.
Dès que la main est suffisamment proche de l'écran, un symbole spécifique, donne sa position courante.

MULTIMODALITE DU DIALOGUE

De nombreuses études ont été effectuées sur l'usage comparé de dispositifs de commande (clavier, souris, manche, commande vocale...) (1,2,3,4) mais peu portent sur leur combinaison. D'autres études (5), par ailleurs, ont proposé des structures de dialogue, mais sans données expérimentales quantitatives.

Expérimentation

Nous avons mené une expérimentation (6) dont l'objectif est de comparer entre elles quatre modalités différentes de la même commande. Cette commande consiste à désigner une figuration dans une image de type avionique.
Pour chaque modalité, l'opérateur accomplit un scénario simulant l'acquittement d'une alarme à bord d'un avion de transport. Cette tâche est constitué de cinq sous-tâches consécutives de désignation sur image.

Les quatre modalités sont:
Modalité 1 (voix seule): messages de trois mots pour

indiquer l'action de désigner et la figuration concernée.
Modalité 2 (main + voix): posture manuelle "*désigne*" sur la figuration, suivie de la validation orale "*o.k.*"
Modalité 3 (oeil + voix): désignation au regard suivie d'une validation orale.
Modalité 4 (oeil + main): désignation au regard suivie de la validation manuelle "*o.k.*"

Résultats

La figure 2 donne les temps de réponses pour chacune des cinq désignations en fonction de la modalité employée. Ces temps de réponse sont moyennés sur 7 sujets, effectuant chacun 6 fois la séquence suivant chacune des modalités. Les segments verticaux donnent l'intervalle de confiance à 95%. L'analyse statistique montre que les facteurs sujet, modalité et rang de répétition ont un effet significatif ($p<0,001$) sur le temps d'exécution total du scénario ; elle montre également que les modalités 1 et 4 d'une part, ainsi que 2 et 3 d'autre part constituent deux groupes statistiquement homogènes.

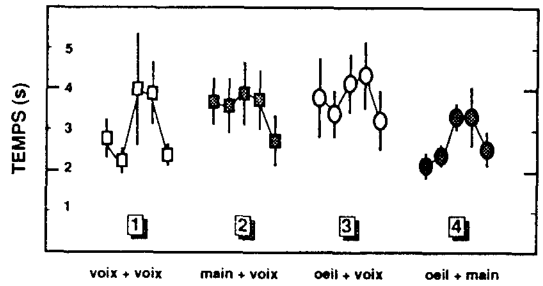


Figure 2: Temps moyen d'exécution des cinq sous-tâches pour quatre modalités de dialogue et intervalles de confiance à 95%

Discussion

Les modalités 1 et 4 ont des temps moyens d'exécution plus courts que celui des modalités 2 et 3.
Les modalités 1 et 4 , par opposition aux deux autres modalités, font appel à des ressources issues d'un domaine unique (7): le domaine verbal pour la modalité 1 (voix seule), et le domaine spatial pour la modalité 4 (oeil/main). Leur performance propre dans l'exécution des deux premières tâches et de la dernière tâche est notablement meilleure que celle des modalités 2 (main/voix) et 3 (oeil/voix) qui, elles, font appel à des ressources à la fois verbales et spatiales.

Cette différence se retrouve également au niveau du temps d'apprentissage. En effet, l'étude de l'évolution du temps d'exécution de la tâche globale, en fonction de son rang de répétition a révélé un effet d'apprentissage plus important pour les modalités 2 et 3 que pour les deux autres.

Toutefois, cette disparité de performances entre modalités mérite d'être relativisée dans la mesure où elle n'est pas observée pour la totalité des sous-tâches: le temps d'exécution des troisième et quatrième sous-tâches du scénario est identique d'une modalité à l'autre. Une des spécificités de ces deux sous-tâches au regard trois autres réside dans le caractère aléatoire de la localisation et la dénomination des symboles à désigner dans l'image.

De plus, l'étude statistique du temps d'exécution du scénario montre qu'il n'y a plus de différence significative entres modalités après la période d'apprentissage (c'est-à-dire à la sixième répétition du scénario). On gardera toutefois à l'esprit le faible niveau de complexité des tâches expérimentées.

MANIPULATION D'OBJETS VIRTUELS

Une modalité particulière de commande que nous avons mise au point consiste en une manipulation d'objets virtuels: l'opérateur modifie, en temps réel, la représentation d'un objet sur l'écran par des mouvements de la main et des doigts.
En utilisant le lexique gestuel, il peut ainsi manipuler un commutateur rotatif virtuel, déplacer une figuration dans l'image, faire disparaître une figuration (8),...

La richesse potentielle du canal de retour graphique compense ainsi en partie l'absence de perception proprio-kinesthésique ordinairement mise en jeu lors de la manipulation des dispositifs mécaniques de commande; cette perception renseigne l'opérateur notamment sur la forme, la position, la déformation et le mouvement du dispositif manipulé (9).

Expérimentation

L'expérimentation élémentaire que nous avons étudiée consiste à saisir un objet virtuel dans l'image: le sujet avance sa main à moins de 50 centimètres de l'image et la referme face à l'objet à saisir. Cette opération est intégrée dans une tâche de fond qui consiste à déplacer l'objet dans l'image jusqu'à un but prédéfini.
L'objectif est d'évaluer l'influence de la taille et de la position de l'objet sur l'exécution de la préhension.
La performance de cette opération, que nous mesurons par son temps d'exécution, est fonction du temps de réaction de la machine et des règles de manipulation qui y sont implémentées.
Huit sujets effectuent, de la main gauche, quatre sessions consécutives; pour chaque session la taille du carré constituant l'objet est fixe et sa position initiale correspond successivement aux quatre coins de l'image. La répartition de la taille suivant le rang de la session et suivant les sujets est faite selon un carré latin. Chaque session est précédée d'un essai de saisie non enregistré.

Résultats

L'analyse statistique du temps moyen de préhension révèle un effet significatif (au niveau 0,01) lié à la taille, mais pas d'effet lié au sujet ni à la position ; le groupe des essais à 40 - 60 - 80 pixels constitue un groupe statistiquement homogène différent du groupe des essais à 20 pixels (1 pixel = 1 millimètre).

La figure 3 donne le temps d'exécution des préhensions moyennées sur celles qui ont réussi à la première tentative et l'intervalle de confiance à 95%; elle indique aussi leur

proportion sur le total des saisies, il n'y a plus d'effet lié à la taille.

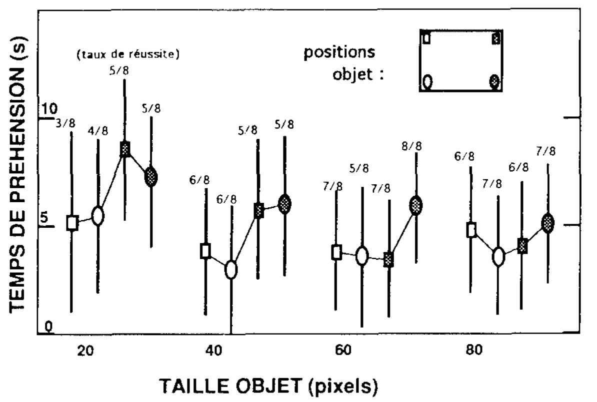


Figure 3: Temps moyen de préhension d'un objet graphique en fonction de sa taille et de sa position sur l'écran; intervalle de confiance à 95% et proportion des préhensions réussies au premier essai.

Interprétation

Le temps moyen de préhension subit une augmentation notable dès que la taille de l'objet est plus petite que 40 millimètres. Cette augmentation est la conséquence de la forte proportion de saisies à tentatives multiples pour les objets de dimensions faibles (le sujet constate que la saisie n'a pas réussi et enchaîne immédiatement par une deuxième tentative sur le même objet, l'exécution correspond à l'ensemble de ces deux tentatives). En revanche le temps d'exécution des préhension réussies au premier essai n'est pas pondéré par la taille de l'objet à saisir, contrairement à ce que prédit la loi de Fitts (10). Celle-ci porte sur des manipulations réelles , par opposition à notre manipulation qui est virtuelle et dont l'exécution est dépendante des temps de réaction propre du Grand Ecran Interactif.

Par ailleurs, la taille de l'objet n'est pas seule responsable de l'existence de tentatives infructueuses; en effet la proportion moyenne de tentatives réussies au premier essai plafonne à 0,81 pour les deux plus grandes tailles (contre 0,53 pour la plus petite).

On notera également l'influence de la latéralité: les saisies d'objets positionnés à gauche sont toujours en moyenne plus rapides que les saisies des objets positionnés à droite (le sujet devait toujours utiliser la main gauche). Cet effet n'est sensible qu'aux faibles tailles d'objet.

CONCLUSION

Le concept de grand écran interactif propose une nouvelle organisation du dialogue homme-système en utilisant d'autres modalités sensorimotrices que celles des canaux visuel et manuel.

La maquette qui a été développée constitue un outil d'étude. Cet outil doit être prochainement intégré dans une plate-forme ergonomique plus complète destinée à évaluer la validité des différentes modalités de dialogue dans un contexte d'intégration cockpit. Il convient cependant d'analyser soigneusement les interactions entre les différentes modalités de dialogue et reconnaître les paramètres influents pour des tâches élémentaires dimensionnantes. Cette indispensable optimisation des processus de dialogue constitue une étape vers une meilleure utilisation des ressources du cockpit et de l'opérateur, conduisant ainsi à une réduction effective de la charge de travail.

REMERCIEMENTS

Les auteurs tiennent à remercier le Docteur A. Léger pour ses conseils dans la définition des expérimentations et dans la rédaction du présent document.

REFERENCES

(1) "Engineering Data Compendium", edited by K.R. Boff and J.E. Lincoln, Wright Patterson AFB, Ohio, USA, 1988.

(2) Santucci, G., Amalberti, R., Menu, J.-P., Malgouyard, P., "Voix et claviers complexes: nouveaux modes du dialogue homme-machine en aéronautique; 2ème partie: Comparaison de deux systèmes de réponse, la voix et le clavier, à des stimulations spatiales non symboliques", CERMA 83-23 (LCBA), Décembre 1983.

(3) Amalberti, R., Batejat, D., Bataille, M., Nessler, F., Santucci, G.-F., "Etude expérimentale d'un moyen de commande en aéronautique intégrant un clavier à labels variables et une commande vocale", CERMA 85-26 (LCBA), Juillet 1985.

(4) Amalberti, R., Santucci, G.-F., Pradella, Valot, "Codage sémantique et codage verbal (réponses à la voix et au clavier à des stimuli colorés)", CERMA 85-35 (LCBA), Octobre 1985.

(5) Bisson, P., Soulard, P., Raimondo, A., Aknin, P., "Interfaces homme-machine multimodales et adaptatives pour systèmes opérationnels", in "Informatique '92, l'interface des mondes réels et virtuels", mars 92, pp.591-603.

(6) Perbet, J.-N., Favot, J.-J., Barbier, B., "Interactive display concept for the next generation cockpit", in SID 91 Digest, Mai 1991, pp 487-490.

(7) Wickens, C.D., Vidulich, M., "S-C-R compatibility and dual task performance in two complex information processing tasks: threat evaluation and fault diagnosis", University of Illinois at Urabana- Champaign EPL-82-3/ONR-82-3, December 1982.

(8) Perbet, J.-N., Favot, J.-J., Barbier, B., "Système de dialogue multimodal pour cockpits futurs", in "Advanced Aircraft Interfaces: The Machine Side of the Man-Machine Interface", AGARD CP 521, 1992, Paper 20.

(9) Cadoz, C., "Le geste de communication homme/machine, la communication instrumentale", ACROE Grenoble, France, 5 mars 1992, p.6.

(10) Holding, D.H., "Human skills, Second Edition ", John Wiley and Sons Ltd., 1989, p.119.

Immersive Virtual Environments as trainer: system design from a cognitive stance

M. Wierda, P.C. van Wolffelaar, W. van Winsum

Traffic Research Centre
University **of Groningen**
P.O.Box 69
9750 AB Haren
The Netherlands

Summary

Many of today's training-simulators for 'guiding, steering or flying' a vehicle are designed to have a safe, environmentally clean, flexible and cost effective educational environment. It is claimed that the training effectiveness can be increased significantly if the starting point of the design would be shifted from the 'enabling technology' position to a cognitive approach of the task to be learned in the simulator. An outline is given of this approach, encompassing a behavioral task-analysis, a cognitive process model and an analysis of the educational goals in terms of cognitive and perceptual skills. It is concluded that knowledge in the domains of cognitive science and artificial intelligence is hardly used while this knowledge may bring about training simulators of a significantly other quality.

1 Introduction

Raise for yourself the following academic question. Someone asks you to build a training simulator for, say a military combat helicopter. Only the best apparatus is good enough, the simulator must be state of the art. You'll have an unlimited budget, but you'll have only one year. What would you do?

Probably you'll acquire an Avens & Sutherwater or Clearsand Picture graphical computer, the best, with full-screen texture rendering and anti-aliasing capable of rendering a trizilion polygons. Next you might consider building a cockpit but buying a complete one from the helicopter manufacturer is more easy. The cockpit is put into a dome, on an Oilpressure Industries six degrees of freedom motion platform. You'll probably make sure that this company also provides for the software drivers to toss around the simulator in accordance with the (subjective) motions of the Helicopter. On top of the Heli you might mount multiple video-projectors. Or, if you are modern, you might use a Head Mounted Display, allowing to omit the dome, projectors and cockpit. Next step might be hiring software experts who take care of the vehicle's dynamic characteristics by implementing process functions. They also might use Plurigem, a software package designed for 3D modelling, to build a combat environment including the graphical representation of a number of potential targets, for example a tank, nicely tucked away in a brushwood. Luckily for the trainee, the graphical computer allows for an 'infra-red-image' visor. A last step might be a visit to the Heli manufacturer's training centre to find out how the flight training is structured. This knowledge is used by another bunch of software experts who hurry to schedule a flight-training plan, inclusive of the educational goals and required skill levels in each stage. Then, once more in a hurry since almost 12 months have past, you might present the full functional prototype to the client. He is very pleased and orders a dozen simulators. You might feel relieved and satisfied since you have done everything you possibly could have and the Heli-simulator probably has the highest possible trainingseffect.

But have you really done everything you could and does the simulator have an optimal educational effect?

The main argument of this paper will be that a significant improvement is possible and a crucial step is overseen. Or even worse, the starting point of the project might be wrong. This is true if, and it must be stressed that it is a conditional argument, the goals of the use of the simulator are 'only':

1 - cost effective training compared to a real-flight-training, given its independence of weather, logistic requirements, operational costs etcetera;

2 - the safety of the pilot, instructor and persons in the vicinity of the training facilities;

3 - the facility to train 'infrequent' hazardous scenario's;

4 - to compare behavioral improvement over time and between trainees in, possibly, 'exactly' the same flight environment,

and that the goal is NOT to use the simulator technique to qualitatively enhance the learning process in both the operational control of the vehicle and the tactical and

This paper is based on a research program initiated and supported by the Ministry of Transport and Public Works, Directorate-General of Transportation.

Presented at an AGARD Meeting on 'Virtual Interfaces: Research and Applications', October 1993.

strategical skills (In section 4 the three distinguished behavioral levels will be defined).

This chapter will outline a strategy for building a Virtual Training Environment (VTE) aiming at the full-blown education/training of men in control of a vehicle. Starting point of this strategy is locating 'the man at the wheel' (or, generally 'at the vehicle controls') in the heart of system. As such, VTE enabling technology (the capabilities of graphical computers, motion platforms, Head Mounted Displays, force-feedback, audio and headtrackers, see for an overview Wierda, 1993), has no other role than the formation of restrictions. In other words, we will follow the line thought of *what* a trainee should be learned and *how* and then, secondary, find out how it can be achieved by using the 'goodies' of the enabling technology. The latter task will be addressed only very shortly since today's technical restrictions will be outdated tomorrow.

Throughout the text explicit examples of a VTE for driving will be used, which is built at the Traffic Research Centre. However the conclusions of the approach and recommendations should be applicable in the design of other VTE's as well and in particular for training of the skills to manoeuvre through a space (driving, flying airplanes and helicopters, controlling trains and armored vehicles).

The strategy has four stages, each will be outlined in subsequent sections. The first step is describing the required task-performance in elementary behavioral elements. The result is called a normative analysis (section 2). The analysis serves two purposes in the formulation of a cognitive process model of the task, step two in the approach (section 3). Firstly the normative analysis prescribes the required output of the cognitive process model and secondly it allows an assessment of the relevance of each sensory channel for building up an internal, mental model of a particular task-environment. As such the process model concentrates on the way humans form mental representations of an environment and how these representations are used to perceive 'changes' in that environment. In a subsequent stage a formalization of 'learning' is given in terms of changes in the internal, mental model (section 4). In a concluding section some typical aspects using a Head Mounted Display in a VTE are discussed (section 5). Based on the cognitive model and the educational goals a VTE may be explicitly designed to bring about the required internal changes in the trainee.

Requirements for 'a' VTE include what elements of the environment must be present, how feedback should be given and what sensory channels are to be used to generate a situational awareness in the trainee that allows him/her to learn and to guarantee that the effect of training will generalize to the task in the 'real world'. This last step is not dealt with since it will be different for each type of task, only examples of the TRC driving simulator will be given. This chapter is concluded by discussing some critical system components of a VTE, in particular the Head Mounted Display.

2 Analyzing the required behavior

A normative task analysis is a list of necessary behavioral elements for performing a specific task adequately. For convenience and usability the elements are clustered around so called manoeuvres. For instance in driving all finely detailed behavior when exiting a highway is given (McKnight & Adams, 1970) or in riding a bicycle the required behavior is scrutinized around manoeuvres as 'turning left on a non-regulated intersection' (Wierda et al, 1989). The resulting taxonomy is *not* intended to, nor capable of predicting and/or explaining observed human behavior psychologically. In fact the analysis even does not take into account that the task is normally performed by a human. If we would be able to build a fully automated car then the best guarantee that it will drive safely would be that the automaton could generate behavior according to the taxonomy.

A taxonomy may have many purposes, in the strategy of designing a VTE only two will be used. Firstly, it prescribes the range of behavior the cognitive process model should account for, or better, should predict. Secondly, the taxonomy can be used to evaluate the relative significance of each sensory channel (seeing, hearing, olfactories, propriocepsis, kinesthetic etcetera) in the overall perception of a specific task-environment. As such this analysis results in requirements for the cognitive process model with respect to 'perception' and it allows an assessment of the required fidelity of the VTE components. If, for example, one is designing a VTE using a Head Mounted Display (HMD) for a Stinger launching site it may be clear that the operator needs to see the incoming jetfighter in time: the pictorial resolution of the HMD must allow the perception of the 'enemy's jet' from quite a distance (see Jense & Kuijper, 1992)

3. A cognitive process model of the task to be learned

A taxonomy of behavioral elements required for a particular task can be used to find the relevant perceptual goals of the different perceptual channels. In next paragraph an example is given for 'visual perception during driving'. Perceptual processes in other sensory channels are skipped, firstly, because the examples given are from the 'driving' task in which visual perception is dominant (both in the taxonomy and in accident causation, see Wierda, in press, Staughton & Storie, 1977) and, secondly, since an elaboration on all sensory channels would take too much space. It should be noted that the method of analysis might be valid for any task-environment. The example below is an excerpt, for a full version see Wierda and Aasman, 1991.

Examples of *visual perceptual goals* while driving:

- 1 Determination of lateral and longitudinal position, changes in these positions and alternations in the changes (lateral and longitudinal speed and acceleration respectively)
- 2 Same as 1 but for heading angle
- 3 Detection of obstacles
- 4 Localization and reading of route indications
- 5 Localization, classification and recognition of roadusers
- 6 Recognition of prototypical, actual traffic situation
- 7 Control over the orientation of the selective visual perceptual system (via body-, head-, and eye-movements and shifts of attention)

The perceptual cognitive process model must be capable of achieving, at least, the enumerated goals. In the outline of the theory, called 3 1/2D model, we will omit an extensive discussion of the low level visual processes such as contour detection, based on motion of a 'blob' against a background, detection of closed contours by boundary tracing and the detection of separated fields by distinguishable 'features'. We will assume that elementary visual routines are capable of deriving an internal representation from the retinal impression (in other words: 'data driven'). For a description of the full theory see Wierda & Aasman, 1991.

Results of elementary visual routines add to a representation that captures the visual environment in separated blobs while the orientation of the surface of the blob is roughly known.

A qualitatively important step is the transformation from this low level 'blob' representation (no objects and backgrounds are identified yet!) to an object centered spatial representation. The 3 1/2D theory claims that a blob is analyzed in terms of a main and an auxiliary axis. The first roughly indicates the orientation, for example a vector running from the feet to the head in case we perceive a human torso, while the auxiliary axis indicates the 'volume' of the object. These internal representations are called generalized cones or object skeletons, the idea of 'summarizing' objects originates from Marr (1982). Complex objects, for example a human figure inclusive of torso, head, arms, hands legs and feet, are composed of series of generalized cones. The main axes of the constituents are hierarchically connected to the prime main axis via slots. The latter specify the allowed movements of the elements with respect to the main axis. The number of formal variables (axes, degrees of freedom) is surprising low with the consequence that vast numbers of hierarchical constructed objects can be remembered with a minimum of storage capacity. Moreover recognition is simplified tremendously: any bottom up perceived object may be compared with any remembered skeleton, from *any* viewing angle by internally manipulating the remembered skeletons. Examples of the internal manipulations are rotation of the main axis to compensate for the viewing angle and enlargement to compensate for viewing distance.

A skeleton representation and the transformation with the set of axes can explain the recognition of a stationary object. However, even during straight driving only a fraction of the retinal input is unchanged, yet we are capable of recognizing an entire scene in a flash. Wierda and Aasman, 1991, proposed to add a single vector to an object's skeleton representation that indicates the direction and pace of movement along the main axis. This aspect of an object is represented explicitly and, as such, is 'remembered' as an integral part of a three dimensional (3D) object. It allows the recognition of objects, for example a nearly 'invisible' car (sic), by degredated contours *and* its typical speed. For this reason the theory has been called the 3 1/2D theory: the three dimensions of space and an abstracted dimension of time. The long term memorized skeletons are easily updated with values for the axes and vectors when comparable but significantly different objects and situations are encountered, estabilshing a 'working memory' version. Among others the long term effect of learning is that 3 1/2D models of objects are clustered into compositions, forming new 3 1/2D models. It is claimed that infra-structure is represented as a 3 1/2D model as well as objects. Composition of these models together with those for moving objects results in *prototypes* for dynamic, 3 1/2D *situations*, in which spational and temporal relations are explicitly represented. Note that these complex prototypes are *learned* when the models are encountered jointly. An example is the formation of a complex 3 1/2D model for a *typical situation* on an intersection, inclusive of the most likely presence and place of roadusers on collision course, from single 3 1/2D prototypes of 'cars', 'pedestrians' and 'infrastructure' (see Wierda and Aasman, 1991, pages 68-71).

An important aspect of adequate spational/temporal visual prototypes is the use of the prototype's parameter values, acting as default terminal values, when these parameters for axes and speed vectors are not immediately available from bottom up perception. This process is considered one of the most important pathways in 'Top Down' or 'cognitively driven' perception. The activation of default values may explain why we (as 'experts' in a certain task) are capable of *generating* a vivid and detailed awareness of our spatial environment even when visual input may be seriously hampered. In next paragraph the shift from bottom up to top down perception is used to explain expertise.

An important difference between novices and experts in controlling a vehicle is the amount of a priori knowledge, structured in the visual prototypes, that are used in perceiving the environment. The task of the expert may be limited to testing his hypothesis based on the Top Down knowledge from the prototypes while the novice has to extract far more 'knowledge' about his task environment Bottom Up, in other words via his sensory channels. This claim may have great consequences for the designer of a VTE: the virtual environment for an expert must be consistent with his expectations and may need no detailing except for critical elements, in other words 'visual cues', by which prototypes are recognized. In contrast the VTE for a

novice needs to have high fidelity with respect to the sensory stimulation. To put it boldly, a simulator for training novices must represent the task environment in details.

4. Learning in a VTE from a cognitive stance

In this section some recommendations for a VTE will be given that are derived from the 3 1/2D theory and the taxonomy. Before doing so a distinction will have to be made in task levels when 'making a trip or flight' with a vehicle. The goal of distinguishing levels of performance is to stipulate the applicability of the taxonomy (section 2) and 3 1/2D theory (section 3) for a wide range of VTE's. The levels can be discriminated, among others, by placing them on a time dimension: while a control level encompasses tasks ranging from milliseconds to seconds, the manoeuvering level takes seconds to minutes and the strategical level may take hours to years (Michon, 1985). On the *strategical level,* one decides how a trip has to be made: by train, taxi, car or whatever. Once a mode of transportation is chosen and the trip is started only every once and a while a decision about what route to follow must be made. Decisions on the strategical level determine the task-environment on the *manoeuvering or tactical level.* On this second level discrete decisions are made on short term trajectories and actions. Examples are what path to follow when negotiating an intersection with the intention to turn left. Also included are the visual search strategies: for example, one 'recognizes' that a potential dangerous situation may be encountered and starts looking for cues of the dangerous object or person. Driving in a residential area with parked cars and playing children is a practical example.

The *control level* of task performance, the last and 'lowest' level, includes high rate first-, second- and third-order control loops. For instance the control of lateral position requires constant adjustments by steering (first order loop). For this reason the level is also called the operational level. In flight, to give an other example, the control altitude, rate of ascent and descent and changes in these rates are first, second and third order control loops respectively. Control tasks are carried out by human subjects by executing 'automatic action patterns', provided that they are experienced. We may add that these tasks require a finely graded representation of the vicinity. It is important to note that tactical decisions have a direct effect on the operational level: tactics define the operational goals. For example, the recognition of the potential dangerous situation of parked cars and playing children should give new parameters to the control loops: one is inclined to brake faster and harder.

The significance of the distinction in task-levels lies in the fact that qualitatively different learning environments are required for each level and therefor differently designed VTE's. In the following paragraphs a VTE for the control level and maneuvering level will be discussed.

If a pilot (or a driver) needs to be trained in operational control of the plane (or car) using a simulator, the fidelity of the controls *and* their effect must be high, probably requiring a six degrees of freedom motion platform and ergonomically well designed pedals, yokes and other controls. In other words, such a VTE may turn out to be very costly. Above that, the operational control of a vehicle is a typical 'perceptual-motor' skill. Generally the speed of acquiring this type of skill can be described by a power law (Newell & Roosenbloom, 1981). And indeed Wierda, Brookhuis and Van Schagen (1987) found a power law for the speed with which young children learn to control their bicycle. If the finding may be generalized to other types of vehicles we might conclude that the trainee learns the control task quickly during the first hours of experience. After having arrived at a relatively high level of performance in a short time the learning process continues endlessly but at a very slow pace. In this context it makes hardly sense to build an expensive VTE to train subjects on the operational level: the required high fidelity of the controls, vehicle model, visualization hardware and motion system require much effort and a huge budget while 'the real vehicle' might be necessary for a limited amount of time. Yet, the VTE for a helicopter pilot described in the introduction will not be capable of training anything else but the operational control. As such the justification of the use of a simulator in the training of operational control of a certain vehicle is a matter of weighing the 'costs' per hour of the simulator and the real vehicle. Costs in this context need not to be restricted to financial consequences but may also refer to effects on the environment and the safety of the pilot and instructor.

A simulator's flexibility in choosing and designing environments seems to make it the ultimate training device for tasks on the manoeuvering level. However, the design of a VTE to be applied in training on this level is rather complex given the following line of thought. Acquiring expertise in manoeuvering a vehicle is achieved by the formation of hierarchical structured 3 1/2D prototypes, see section 3. As such the trainee must experience a wide range of interactions with others in the environment, whether in combat, driving or flying, while the others behave to a large extend 'naturally'.

Generating 'natural' behavior in real time is truly a complex task since, we will use a driving simulator as an example, the other roaduser will have to interact among each other and with the subject in the simulator. The range of artificial behavior of the 'other roadusers' in the simulator environment should be as large in real traffic, otherwise it would never appear to be natural. The others should overtake, negotiate intersections, make emergency brakes, 'swerve' naturally, slow down for curves etcetera. Perhaps surprisingly a taxonomy and a cognitive model such as described in sections 2 and 3 are badly needed to 'move' the artificial roadusers (See Van Winsum, 1991). Effectively 'other' roadusers in the VTE are autonomous in their interactions, inclu-

sive the interactions with the trainee in the VTE. This would mean that the situations encountered by the trainee are not predictable, once the simulation is started. This would be a major drawback: an instructor, a human or possibly an automaton, needs the capability to bring the trainee in those traffic situations in which the formation of spational/temporal prototypes is optimal. As such the 'other roadusers' must be controllable to a certain extend. Therefore the VTE at the Traffic Research Centre (TRC) has been equipped with a Scenario Specification Language allowing the deliberate set up of precisely defined traffic situations (Van Wolffelaar & Van Winsum, 1992). These scenarios that define the intentions of other roadusers are triggered when the trainee passes a certain geographical point in the driving environment. The trainee will never know that the traffic situation is a 'set up': he will experience his ride through the traffic environment as natural. Currently the VTE is tested on effectiveness (Wierda, 1993).

In conclusion of this section it is stated that a VTE is very promising to learn people to navigate a vehicle in an interactive environment. However an extensive cognitive model of the driver's (pilot's) task is required firstly since it is the base for the Artificial Intelligence of others in the environment and, secondly, since it is required to assess the ongoing formation of spational/temporal prototypes in order to present the trainee exactly those traffic (flight-combat) situations in which the formation of prototypes is maximalized. One must not underestimate the required effort to build and apply these models.

5. Using a Head Mounted display in a VTE

The core of learning how to manoeuvre a vehicle in a crowded environment is developing the skill to apprehend the actions and intentions of 'others' in the environment. Only with this skill one can take the right decisions on what path to follow, what speed to maintain and what formal rules to apply. Since the main source of information about intentions of 'others' is the way they move their vehicle through the environment, the training facility should focus on these movements. Therefore the VTE should have the facility to visualize explicitly the movements of others and the movement of the trainee's vehicle *off line* by which implicitly the intentions and interactions in the situation are made clear to the trainee. In practice, a trainee in the VTE at the Traffic Research Centre will encounter specific scenario's (without knowing it) in which his behavior, inclusive of choice of speed, gear, steering, visual scanning behavior etcetera, is judged. When this behavior is sub-optimal the simulation is postponed and the scenario is played back. The significance of the Head Mounted Display is the capability to look around during the play-back. Above that different viewing angles can be chosen. Currently the viewing point can be set in any of the cars of 'others' in the situation and it may also be set in a helicopter viewmode. In the latter viewing point the three dimensional/temporal properties of the traffic situation become crisp clear. It is important to note that this form of augmented feedback (see Sanders, 1991) is immediate: the intention of the training is to accelerate the spatio-temporal reasoning and therefore the feedback is given in a spatio-temporal form. This is considered an important improvement relative to a classical instruction. In the latter case an instructor gives *verbal* hints, for example "you should pay more attention to motorized traffic coming from the right if you negotiate such an intersection". The trainee must interpret the message, transform it into a spatio-temporal representation and applying it to the just encountered traffic situation while he/she could already be involved in a subsequent traffic situation.

The merits of an HMD in a VTE are manifold, but general. The most important one is the impact of the VTE on the situational awareness of the subject. This effect has been explained by the isolation of the subject's sensory system from the 'real, actual' world. Furthermore the stereoscopic view makes the spaciness of the environment so compelling that the entire environment will become 'believable', even though seeing depth through stereopsis (Marr & Poggio, 1979) is of no importance (in flight) or relative importance (driving) in the actual task (Wierda, in press).

A last to be mentioned merit of an HMD in a VTE designed for training on the tactical level is that the trainee can be prevented to focus his attention on the instruments and vehicle controls: they simply are not visualized. This state of affairs may sound odd, but it is a commonly heard complaint from instructors (both in flight and driving lessons) that trainees spent too much time looking at odometers and other instruments and take too much time looking at their hands when applying the vehicle's controls. By simply not visualizing the instruments and controls the trainee is forced 'to look outside' and forced to build up rapidly an internal representation of the spational layout of the controls.

6. Conclusion

Virtual Training Environments using a Head Mounted Display are an important extension to simulator technology. An HMD-VTE combines the capabilities of simulators and the compellingness through immersion of the HMD technique. However, it is observed that even large scale projects to build VTE's for flight, driving and other vehicle control are technology driven, while we expect the largest improvement in training effectiveness on the manoeuvering level of the task. As such complex cognitive process models are required for two reasons. Firstly, the training environment should focus on the interactions between 'others' in the environment, whether other roadusers in driving or enemies in a combat flight. As such these 'others' must behave 'real time' and ecologically valid. Secondly, if specific environments are to be presented to the trainee to optimalize the training's effectiveness one must understand

the learning process of the trainee in terms of the formation of spatio-temporal prototypes. It is claimed that the disciplines of cognitive science and Artificial Intelligence can provide us with the required process models, in fact this has been partly done in the case of the VTE at the Traffic Research Centre. However one must not underestimate the required effort to apply these models and therefore it would be wise to take the cognitive approach as a starting point in the design of a VTE instead of starting to build a VTE and, once it has been realized in hardware, being forced to conclude that the VTE might be great for training the control of a vehicle on the operational level but that the tasks on the tactical level must be excluded.

References

Jense, G.J. & Kuijper, F. (1992). Virtual reality research at TNO/FEL. *Proceedings of the 2nd ESA Workshop on Simulators for European Space Programmes* Report ESA-WPP-039. Noordwijk: ESA.

Marr, D., Poggio, T. (1979). A computational theory of human stereo vision. *Proceedings of the Royal Society of London*

Marr, D. (1982). *Vision*. San Francisco: Freeman.

McKnight, A.J. & Adams, B.B. (1970). *Driver education task analysis. Volume I: Task Descriptions*. Report 11-7336. Alexandria, VA: Human Resource Research Organization.

Michon, J.A. (1985). A critical view of driver behavior models: what do we know, what should we do. In L. Evans en R.C. Schwing (Eds.). *Human behavior and traffic safety* New York: Plenum.

Newell, A.N. & Roosenbloom, P. (1981). Mechanisms of skill acquisition and the law of practice. In J.R. Anderson (Ed.): *Cognitive skills and their acquisition*. Hillsdale, New York: Erlbaum.

Newell, A.N. (1990). *Unified theories of cognition* London: Harvard University Press.

Sanders, A.F. (1991). Simulation as a tool in the measurement of human performance. *Ergonomics*, 34, 995-1024.

Staughton, G.C., Storie, V.J. (1977). *Methodology of an in-depth accident investigation survey*. Report LR 762. Crowthorne: Transportation and Road Research Laboratory.

Wierda, M., Schagen, I.N.L.G. van, Brookhuis, K.A., Rothengatter, J.A. (1989). *Normatieve analyse van de fietstaak*. (A normative analysis of riding a bicycle). VK-89-29. Haren: Traffic Research Centre.

Wierda, M., Brookhuis, K.A., Schagen, I.N.L.G. van (1987). *Elementaire fietsvaardigheden en mentale belasting*. (Basic skills for riding a bicycle and mental effort) VK-87-08. Haren: Traffic Research Centre.

Wierda, M., & Aasman, J. (1991) *Seeing and driving, computation, algorithms and implementation* VK 91-06. Haren: Verkeerskundig Studiecentrum.

Wierda, M. (1993). *PACT, een geautomatiseerde rij-instructie in een Virtuele Omgeving*. (PACT, an automated driving instruction in a Virtual Environment) VK 93-02. Haren. Traffic Research Centre.

Wierda, M. (In Press). Beyond the eye: cognitive factors in drivers' visual perception. In: A. Gale (Ed.): *Proceedings of the Fifth international conference on Vision in Vehicles*. London: Taylor & Francis.

Winsum, W. van. (1991). *Cognitive and normative analysis of car driving* Haren: Traffic Research Centre.

Wolffelaar, P.V. van & Winsum, W. van. (1992). *A new driving simulator including an interactive, intelligent traffic environment* Proceedings of the 3rd International Conference on Vehicle Navigation and Information Systems. Piscataway, New Jersey: Institute of Electrical and Electronic Engineers.

Creation of a Virtual World to Study Human Spatial Perception during Sustained Acceleration

Tamara L. Chelette and Robert L. Esken
Armstrong Laboratory, AL/CFBS Bldg 29, 2245 Monahan Way
Wright-Patterson Air Force Base, OH, 45433-7008, USA

Eric J. Martin
Systems Research Laboratories, Inc.
2800 Indian Ripple Road
Dayton, OH, 45440, USA

1. SUMMARY

The staff of the Combined Stress Branch has completed the integration of a system to allow quantitative measurement of perceived attitude while under sustained acceleration. Equipment involved included the computer control system of the Dynamic Environment Simulator (DES), a computer generated graphics system, a virtual world helmet mounted display, and a tactile device for reporting attitude perception. The use of a new perceived attitude measurement system in this experiment required not only the technical achievement of the distributed system on the DES, but also required a battery of parameter characterization and basic psychophysical performance studies. In addition, we recorded several confounds and issues concerning the use of a helmet mounted visual system for attitude information as well as head and neck support limitations of such a system. Experimental results include basic psychophysical accuracy and precision, evidence supporting the haptic system sensitivity to a G-excess illusion (even while the vestibular system is maintained at a constant position relative to the G vector), and modeling of pooled response that supports and quantifies the vestibular component of the G-excess illusion.

2. INTRODUCTION

Spatial Disorientation (SD) of pilots continues to be a very serious human factors issue in the United States Air Force and Navy [1,2,3]. In the Air Force, SD results in 8 to 10 aircraft crashes and pilot deaths yearly [4]. A common mishap scenario is low level banked turns while looking up at a lead aircraft [5,6]. Many scholars have theorized and/or investigated the human vestibular response to tilt while varying the magnitude of the gravitoinertial field. Perceptual response metrics have included [7] placement of some external object, simulated aircraft controlled flight recovery [8,9], verbal descriptions of scales [10] or vection [11], manual keyboard inputs [12], and coded hand signals [13]. Some researchers have combined metrics in order to overcome the limitations or assumptions required of a single response [12]. Placement of external lines or points can be influenced by optical illusions [14]. Control recovery metrics bring the subject into an interactive role that immediately influences perception [15]. Subjective reporting of perception usually requires subjects to translate an internal sensation to some other medium such as words or force.

In addition to the difficulty in objectively measuring vestibular response, ground based research is confounded by the necessity to use high angular velocities to alter the gravitoinertial field strength. Thus, vestibular response is a product of both tilt and angular velocity.

Although there is some evidence that the semicircular canals can become tilt sensitive under the influence of alcohol [16], most of the cited studies support the theory that the otolith organs are the primary vestibular sensors of tilt, and many of them support the theory that shear displacement between macula and otoconia is the excitation stimulus [17, 18]. The G-excess illusion is believed to originate in the otoliths. While in a prolonged coordinated turn, pilots often must look out the cockpit to find other aircraft or survey a target. If the head is tilted with respect to the aircraft, and the aircraft is sustaining greater than 1 G_z caused by the banked turn, an illusion of excessive head tilt may result giving rise to the interpretation that the aircraft has rolled out of the turn to some extent (Figure 1). If a "correction" is made for this erroneous sensation, the aircraft can overbank and loose altitude. If this situation persists, the aircraft can slice downward at a fatal velocity. The G-excess effect has been implicated in over 40 of the 70+ SD related USAF Class A mishaps since 1982 [19]. It is this particular illusion, the G-excess effect, that was investigated in this series of experiments.

The objective of this study was to determine if the effect of head tilt in a greater than one G environment on perception of attitude could be demonstrated and quantified using a ground based human centrifuge. Head tilts were accomplished not only in the body pitch axis, but also in body roll axis via head yaw. Although this type of head movement is fairly unique in vestibular research, it is a common and necessary head position in an aircraft cockpit. For example, formation flying

Presented at an AGARD Meeting on 'Virtual Interfaces: Research and Applications', October 1993.

may require a pilot to maintain a gaze up and toward one shoulder by 45°. During air to ground missions, pilots may yaw their heads as much as 120° and gaze downward to assess munitions targets. Air to air combat can frequently require the 'check six' maneuver where the pilot must attempt to look directly behind him ('six-o'clock'). The focus of the response measurement was on perception of self attitude as reported by hand position. This approach was designed to take advantage of the intuitive behavior most people exhibit of placing the palm of their hand parallel to the surface of the earth when asked to report the horizon. This unique combination of stimulus-response was designed to be sensitive to the potential illusions elicited on a ground based centrifuge.

3. METHODS

The general method for this experiment was to collect a measure of the subject's perceived orientation while s/he was at a steady state G level and actively accomplishing some known head tilt. The greater than one G was provided by a man-rated centrifuge and the head aiming task was accomplished with a visual virtual reality system. The subject's perceived orientation was collected by instructing them to orient the palm of their right hand such that it was parallel with the perceived horizon while their hand was suspended in a multi-gimbaled transducer. The details regarding this equipment as well as the motivation and methods of the experimentation follow.

Equipment

The primary piece of equipment was the Dynamic Environment Simulator (DES), a three axes, 19 foot radius, man-rated centrifuge located at the Armstrong Laboratory, Wright-Patterson Air Force Base, Ohio. The acceleration was imposed by the rotation of the DES and "auto-vectoring" of the cab such that the resultant force vector acted along the longitudinal axis of the body (G_z). Any cab pitch or roll deviations introduced as independent variables in this study were with respect to this resultant G_z vector. In accomplishing cab tilts, the entire subject environment including the subject's seat is tilted. Subjects were seated in a F-16-like 30° back seat, restrained with a five point harness, and provided cardiovascular support with a CSU-13B/P standard anti G-suit. Subjects also wore a standard issue HGU-55/P flight helmet and a MBU-12/P oxygen mask, primarily to support communications and the visual virtual reality system.

The method for incorporating the head aiming task involved a visual virtual reality system and an associated head tracker (Figure 2). The visual virtual reality system, developed by VPL Research, Inc., was an Eye Phone model number EP-01 driven by an XTAR Corp. Super Falcon 4000 graphics package with a 30 Hz update rate. The head tracker used was a Polhemus Navigation Sciences 3Space Isotrak system utilizing low frequency magnetic field technology to determine the position and orientation of a six degree-of-freedom sensor. The sensor was mounted on the helmet and the visual scene projected in the eye phones was slaved to the motion of the sensor, and therefore, the motion of the head.

Collection of the subject's perceived orientation was accomplished using a device developed in-house and known as the Tactile Perceived Attitude Transducer (TPAT) [20]. This device consists of an aluminum hand plate with a glove suspended on the underside (Figure 2). When the subject's hand is inserted into the glove, finger and wrist restraints secure the hand such that the back of the subject's hand is firmly affixed to the underside of the hand plate. The hand plate is mounted on a pitch axis gimbal that is anchored to a steel captured bearing in the roll axis. The captured bearing is mounted on a steel pivot providing yaw motion. Movements of the gimbal, bearing, and pivot may be accomplished simultaneously allowing hand motion in three angular degrees of freedom. These movements are detected by potentiometers which are recorded as pitch, roll, and yaw positions. This device takes advantage of the natural inclination to describe one's orientation in space by positioning the suspended palm parallel to the perceived horizon.

Subjects

The nine volunteer subjects (seven male, two female) in this study were all members of the Armstrong Laboratory Sustained Acceleration Panel. They were active duty military personnel with extensive centrifuge experience but limited flight experience. The research protocol and procedures were reviewed and approved by the Armstrong Laboratory Human Use Review Committee.

Target Tracking Task

A visual image was provided by the virtual reality goggles. The image consisted of two components: a spherical target that was driven in software to drift to prescribed locations in the visual field, and a reticule slaved to the position sensor mounted on the helmet (Figure 3). Subjects were instructed to perform as if they were flying at night and watching the moon while simultaneously placing their right hand parallel with the horizon. The head tilts were realized by drifting the spherical target to the desired pitch and yaw angle while the subject followed it with the reticule by moving his or her head as if the reticule was etched on a pair of glasses. Target drift was followed by a twelve second vestibular stabilization period, throughout which the reticule would blink. Upon cessation of the blinking, the subject would orient their hand in the manner described and mark that particular hand position by pulling a trigger on the flight stick with the left hand.

Head Tilt Calibration

Subjects were seated in a 30° back seat and secured with a

five point seat harness. They then donned the helmet and acquired a snug fit. The virtual reality goggles were then mounted on the helmet with a series of nylon straps and a set of spring loaded ear pieces fashioned from a pair of welding goggles. This mounting scheme was designed to decrease the torque resulting from the protruding goggles, which weighed 2 lbs. and 1 ounce, by distributing the load nearer the helmet's center of gravity. Once a comfortable yet firm fit was achieved, the calibration routine was initiated.

The calibration procedure involved slaving the Polhemus position sensor mounted on the top of the helmet and the visual image displayed in the goggles to the head position. This was accomplished by instructing the subject to assume a neutral head position with negligible pitch or roll. The sensor/goggle system was then zeroed, or boresighted, at this position. Using a visible red light pointer mounted on the helmet, the point of light was projected onto a white measuring tape hanging plumb in the centrifuge cab. Distance between the subject's eye and the hanging tape was obtained via a yardstick with a bubble level secured to it to ensure a measurement parallel to the true horizon. The target would then drift to its new position in the pitch axis, the subject would track it by moving his head, and the change in the position of the point of light would be measured. The arcsine of the ratio of these distances yielded the angles accomplished within 1°. These four pitch angles were stored with the data set for each data collection session and used as the actual head position in the data analysis.

Preliminary Experimentation

Since this combination of equipment represented a new technique for perceptual research, extensive preliminary testing was performed to examine issues such as training effects, visual feedback, numerical feedback, accuracy, precision, and non-vestibular sensitivity to tilt under G_z. Descriptions of this testing are beyond the scope of this article but can be found in reference [21]. The following is a brief synopsis of the salient findings.

Spatial information regarding reported position fed back visually confused the subjects when displayed at positions other than straight ahead and thus decreased their accuracy at reporting environmental tilt. Numeric feedback regarding perceptual error did not confuse subjects and thus improved their performance, however no residual group training effect was demonstrated after the feedback was removed.

Repeated measures of reported environmental tilt showed no statistically significant bias in either pitch or roll axis with the head level or pitched upward 45°.

Before designing the final experiment, the investigators wanted to know if the non-vestibular components of perception were also sensitive to an increased G_z environment. When the environment was tilted to counter the head movement such that the vestibular system input remained unchanged, subjects could still accurately report environmental tilt at one G_z. However subjects showed an overestimation of tilt at 3 G_z while in pitch prone positions. Thus haptic input alone is sufficient to cause a pitch illusion.

Final Experimental Design

The experimental design incorporated four cab environment pitch angles (-5°, 0°, 5°, and 10°) and four head pitch angles (-30°, 0°, 30°, and 45°). In order to reduce the number of trials a subject would endure in a data collection session, either cab environment angle or head pitch was equal to zero, resulting in a total of seven paired cab pitch/head pitch permutations. These paired conditions were presented randomly such that the subject never knew whether or not the cab was offset from zero during a given trial. Separate models were constructed with respect to the head pitch and cab environment pitch factors. The cab environment angles chosen were comparable to the expected magnitude of the illusion.

The G_z levels incorporated were earth-normal 1.0, 1.4, 2.0, and 4.0. Four G_z was selected as the maximum G_z level as subjects had difficulty supporting the VPL helmet mounted system at any greater G_z level for the required length of time. The 1.0 G_z trials were accomplished first before proceeding to the induced G_z trials (1.4, 2.0, and 4.0). The induced G_z trials were initially presented randomly until it was suspected that the frequent decelerations associated with a random presentation introduced nausea in more susceptible participants. In order to reduce the number of potentially nauseogenic decelerations, an ordinal presentation was employed (1.4, 2.0, 4.0, 1.4, 2.0, 4.0, etc.) until it was determined that head motion during the tracking task *combined* with a deceleration was causing the discomfort. Thereafter, subjects were instructed to maintain their current head position until the centrifuge arm speed stabilized, then acquire the target and wait the aforementioned 12 seconds to indicate their perceived attitude. The twelve second stabilization period was selected in order to mitigate the dynamics of the semicircular canals from contributing to the perception and reporting of attitude. With this technique, it was possible to return to the random presentation of G_z level. Each subject, therefore, experienced one data collection session with an ordinal G_z presentation and one with a random sequence, except our most susceptible subject, who received the ordinal sequence in both sessions.

Finally, head yaw is a highly relevant movement in the fighter cockpit and serves to translate head pitch sensation into aircraft roll. Thus three head yaw conditions were introduced: 0°, 45°, and 90°. This was accomplished by mounting the

seat at angles of 0°, 45°, and 90° from radial and instructing the subject to look over his/her right shoulder to the radial position. In this way, the pitch axis of the head was maintained in the same plane as the cab axis so that an illusory tilt would be sensed in the same plane as an actual cab tilt. Since seat positions (head yaw) could not be presented randomly from a practical standpoint, all subjects completed their sessions in the 0° seat position first (seat mounted radially), followed by the 90° yaw position (seat mounted tangentially, head pointed radially), and concluded with the oblique 45° seat position (head pointed radially). These 6 exposures (2 per head yaw condition) were accomplished on 6 separate visits to the laboratory.

Cumulatively, the seven paired environment cab pitch/head pitch combinations, the four G_z levels, and the three seat positions (head yaw conditions) resulted in 84 combinations of independent variables. Each of these was repeated twice by each of the nine subjects. Data were recorded in both the pitch and roll axes resulting in over 3000 data points.

4. RESULTS

Four multiple regression models were built using data from the following conditions:

- Reported pitch data when head tilted.
- Reported roll data when head tilted.
- Reported pitch data when cab is tilted.
- Reported roll data when cab tilted.

Inspection of the data revealed that head yaw was an important factor independent of head pitch as it had a profound effect on perceived and reported attitude. This was evident in the conditions of zero head pitch and one G_z where there is an effect of head yaw alone. The effect appears to be nonlinear, thus modeling was tested for head yaw and the square of head yaw.

Inspection of the data with respect to G_z levels reveals a second effect independent of the other treatment effects. This effect was suspected to be due to the centrifuge arm speed. In each model, a term was added that reflected either total G_z, radial acceleration $(G^2 - 1)^{.5}$, or angular velocity $(G^2 - 1)^{.25}$ and the best fit was selected.

The third effect for which the data were tested was a correlation to a term proportional to the difference between the sine of the head pitch (or the cab pitch) and the sine of the neck pitch magnified by some function of G_z level as translated into the pitch or roll axis of the body by head yaw. This is the proposed G-excess effect. Four terms were examined; linear, linear with 30° offset (due to otolith anatomy), nonlinear, and, linear with saturation. A G excess illusion is believed to be the excess tilt sensed beyond that accounted for by the tilt of the neck. Therefore, in predicting the magnitude of the illusion we assume the individual has self knowledge of neck tilt and this must be subtracted from the sensed tilt. This is represented in the equations below as $G^{.25}$ x sin[head pitch] - 1 x sin[head pitch] where the second term is self knowledge of neck tilt. This was the term used to get the coefficients in Table 1 and the equations above.

Pitch Axis				
	Effect	Term	Coeff	p Value
Head Tilted		Intercept	-0.8767	.1868
	1	G	1.4838	.0001*
	2	HY	0.0214	.0064*
	3	$(G^{.25}-1)$xsin(HP)	0.1491	.0193*
	Variance	accounted for		56.9%
Cab Tilted		Intercept	-0.4044	.5968
	1	G	1.4341	.0001*
	2	sin (CP)	0.9175	.0001*
	3	HY	0.0191	.0301*
	Variance	accounted for		79.8%

Roll Axis				
	Effect	Term	Coeff	p Value
Head Tilted		Intercept	-4.3188	.0001*
	1	HY^2	-0.0013	.0001*
	2	$(G^{.25}-1)$xsin(HP)	0.3397	.0001*
	3	$(G^2 - 1)^{.25}$	-2.1135	.0001*
	Variance	accounted for		92.3%
Cab Tilted		Intercept	-4.0421	.0001*
	1	sin (CP)	1.2438	.0001*
	2	HY^2	-0.0013	.0001*
	3	$(G^2 - 1)^{.25}$	-2.2183	.0001*
	Variance			93.9%

Table 1 - Selected terms and significance of each effect in each model (* indicates statistical significance, p<.05).

The best fit term was selected for each effect. Table 1 shows the order of inclusion, coefficients, and resulting significance of each term

The coefficients given below are taken from effect 3 in the top portion of Table 1. These data support the following G-excess illusion magnitudes (in °) as a function of head pitch (in °), head yaw (in °), and G_z level (earth g units):

Mag of roll illusion =
0.3397 x arcsin{$(G^{.25}$ - 1) x sin [head pitch] x sin[head yaw]}

Mag of pitch illus =
0.1491 x arcsin{$(G^{.25}$ - 1) x sin[head pitch] x cos[head yaw]}

When the head is kept level, but the cab (representing the aircraft) actually tilts with respect to the net gravitoinertial vector (uncoordinated turn), these data indicate that subjects will accurately assess the tilt without a significant illusion. This is most likely due to the sensitivity of the other haptic sensing cues that provide information about tilt when actual tilt occurs. This is consistent with the conclusions of the preliminary phases of this experiment.

The nonlinear term was selected for inclusion in the head tilt models because it consistently fit the data better than the other three proposed terms for the G-excess effect. However, none of the differences among the terms was statistically significant. Thus any one of these functions could be used as an estimate. The apparent success of the nonlinear term is likely due to the nonlinear elastic properties of the macula-otoconia interface of the otolith organs. The apparent insignificance of the anatomical otolith orientation (30 ° offset) is likely due to the lifelong adaptation to such. The apparent attenuation of the sine function while under G_z (i.e., the 0.3397 coefficient) is most likely due to the amplified haptic signals and acute attention to orientation during experimental G_z exposure.

5. CONCLUSIONS

The data described above support the hypothesis that pitching the head while in an excess G_z environment (>1) can cause an illusory sensation of vehicle tilt. This illusion occurs in the pitch axis if the head is forward, however translates to the roll axis as the head is turned toward one shoulder. This effect can be reproduced on a ground based centrifuge provided confounding factors are taken into account in the model. Subjects demonstrated that true vehicle tilt up to 10° is accurately assessed without any significant illusion while head tilts in the -30° to +45° range up to 4 G_z can cause illusionary tilts up to approximately 10° as well.

The magnitudes of illusions demonstrated in this experiment were based on a steady state response to a sustained head position and G_z level. However, physiological evidence of rate sensitive otolithic cells [22] combined with in-flight evidence that supports sensitivity to rate of head movement [9] necessitates the caveat that actual occurrence of the G-excess illusion may result in significantly larger transient illusory angles.

6. RECOMMENDATIONS

Pilots should be made aware of the possibility and magnitude of the G-excess effect. Training protocols should include the caveat that head pitches can cause erroneous sensations of under or overbanking of their aircraft. Special attention must be paid at low altitude to avoid disaster [5]. Specifically, an upward head pitch combined with a head yaw into a turn, as is common in formation flying, can result in a sensation of aircraft underbank. Intended corrective action actually overbanks the aircraft, causing loss of altitude. Downward head pitches during turning, as is common during bombing or strafing runs, can cause a sensation of overbanking. Intended corrective action actually underbanks the aircraft, causing altitude gain which could lead to midair collision when in formation flight. Although this illusion can be accounted for in ground based centrifuge simulator testing, pilot training should be provided in flight, not only to avoid confounding sensations, but to teach the necessary flight control behavior in such situations.

Acknowledgments

The authors wish to acknowledge the valuable experimental design contributions of Dr. Alan Benson, Dr. Donald Parker, Dr. Kent Gillingham, and Dr. Henning von Gierke. The authors wish to extend their appreciation also to William Albery, Thomas Hundt, Deepa Naishadham, Chuck Goodyear, Systems Research Laboratories, Inc. and the entire support staff of the Dynamic Environment Simulator Faciltiy. The authors are also grateful to the subjects who patiently participated in these experiments.

7. REFERENCES

1. DeHart R. Human factors problems in the Tactical Air Command. Aviat Space & Environ Med. 1986; 57:725-8.

2. Wanstall B. Helping combat pilots survive. Interavia Aerospace Review. Mar 1990.

3. Bellenkes A, Bason R, Yacavone DW. Spatial disorientation in Naval aviation mishaps: A review of class A incidents from 1980 through 1989. Aviat Space & Environ Med. 1992; 63:128-131.

4. Lyons TJ, Freeman JE. Spatial Disorientation (SD) mishaps in the U. S. Air Force. Paper presented at the 61st Annual Scientific Meeting of the Aerospace Medical Association. Abstract # 62. May 1990. New Orleans, LA

5. Lyons TJ, Gillingham KK, Thomae CA, Albery WB. Low-level turning and looking mishaps. Flying Safety. Oct 1990.

6. English AL. A lifesaver called GCAS. Flying Safety. Jan 1990.

7. Working Group on Evaluation of Tests for Vestibular Function, Committee on Hearing, Bioacoustics, and Biomechanics (CHABA), Commission on Behavioral and Social Sciences and Education, National Research Council. Evaluation of Tests for Vestibular Function.

Aviat Space & Environ Med. 1992; 63:2II(a1-a34).

8. Lyons TJ, Gillingham KK, Teas DC, Ercoline WR, Oakley C. The effects of acoustic orientation cues on instrument flight performance in a flight simulator. Aviat Space & Environ Med. 1990; 61:699-706.

9. Guedry FE, Rupert AH. Steady state and transient G-excess effects. Aviat Space & Environ Med. 1991; 62:252-3.

10. Fulgham, D. Spatial Disorientation. Technology Today. 1989. March.

11. Previc FH, Mullen TJ. A comparison of the latencies of visually induced postural change and self-motion perception. J of Vestibular Research. 1991; 1:317-23.

12. Hosman RJAW, van der Vaart JC. Motion perception and vehicle control. In: Warren R & Wertheim AH, eds. Perception and control of self motion. Hillsdale NJ: Lawrence Erlbaum Assoc. 1990; 399-423

13. Malcolm R, Melvill Jones G. Erroneous perception of vertical motion by humans seated in the upright position. Acta Otolyng 1974; 77:274-283.

14. Cohen MM. Elevator illusion: Influences of otolith organ activity and neck proprioception. Perception & Psychophysics. 1973; 14:401-406.

15. Stoffregen TA, Riccio GE. An ecological theory of orientation and the vestibular system. Psychological Review. 1988; 95:3-14.

16. Money KE. Disorientation with alcohol and acceleration. Presented at the 6th Interservice/Industry Acceleration Colloquium. 1992. Dayton, OH.

17. Grant JW, Best W. Otolith organ mechanics: lumped parameter model and dynamic response. Aviat Space & Environ Med. 1987; 58:970-976.

18. Grant JW, Cotton JR. A model for otolith dynamic response with a viscoelastic gel layer. J of Vestibular Research. 1990; 1:139-52.

19. Freeman, J., (1990), Personal letter from Air Force Inspection and Safety Center toT. L. Chelette re: Spatial Disorientation Mishaps, dtd 11 May, 1990.

20. Chelette TL. Developing a Technique for measuring the G excess illusion. Human Factors Society Test & Eval Tech Grp Newsletter, 1991; 6(3):2-6.

21. Chelette TL, Martin EJ, Albery WB. The Nature of the G Excess Illusion. Technical Report. AL-TR-92-0182.

22. Anderson JH, Blanks RHI, Precht W. Response characteristics of semicircular canal and otolith systems in cats. I. Dynamic responses of primary vestibular fibers. Exp Brain Research. 1978; 32:400-21.

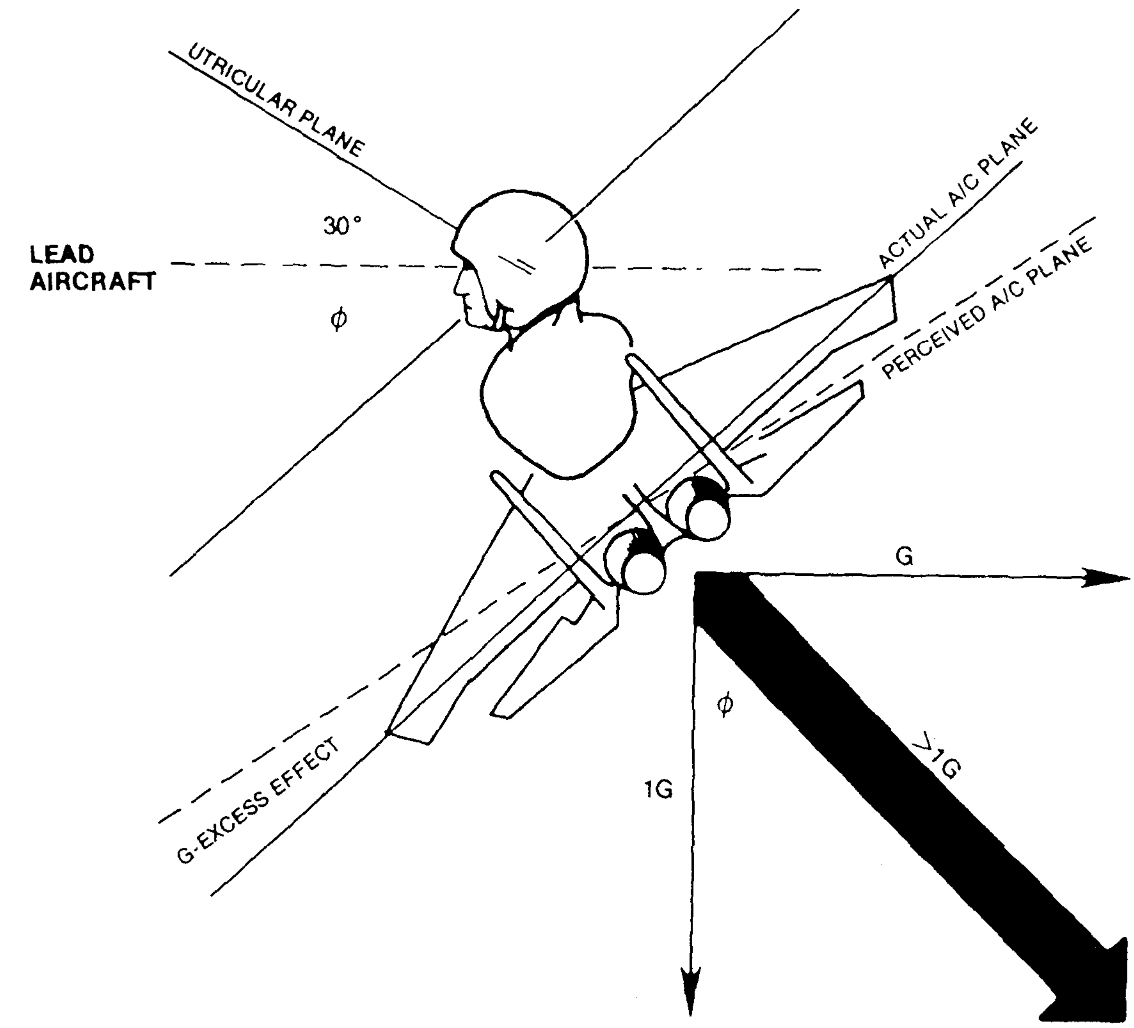


Figure 1. G-Excess Effect and Formation Flying.

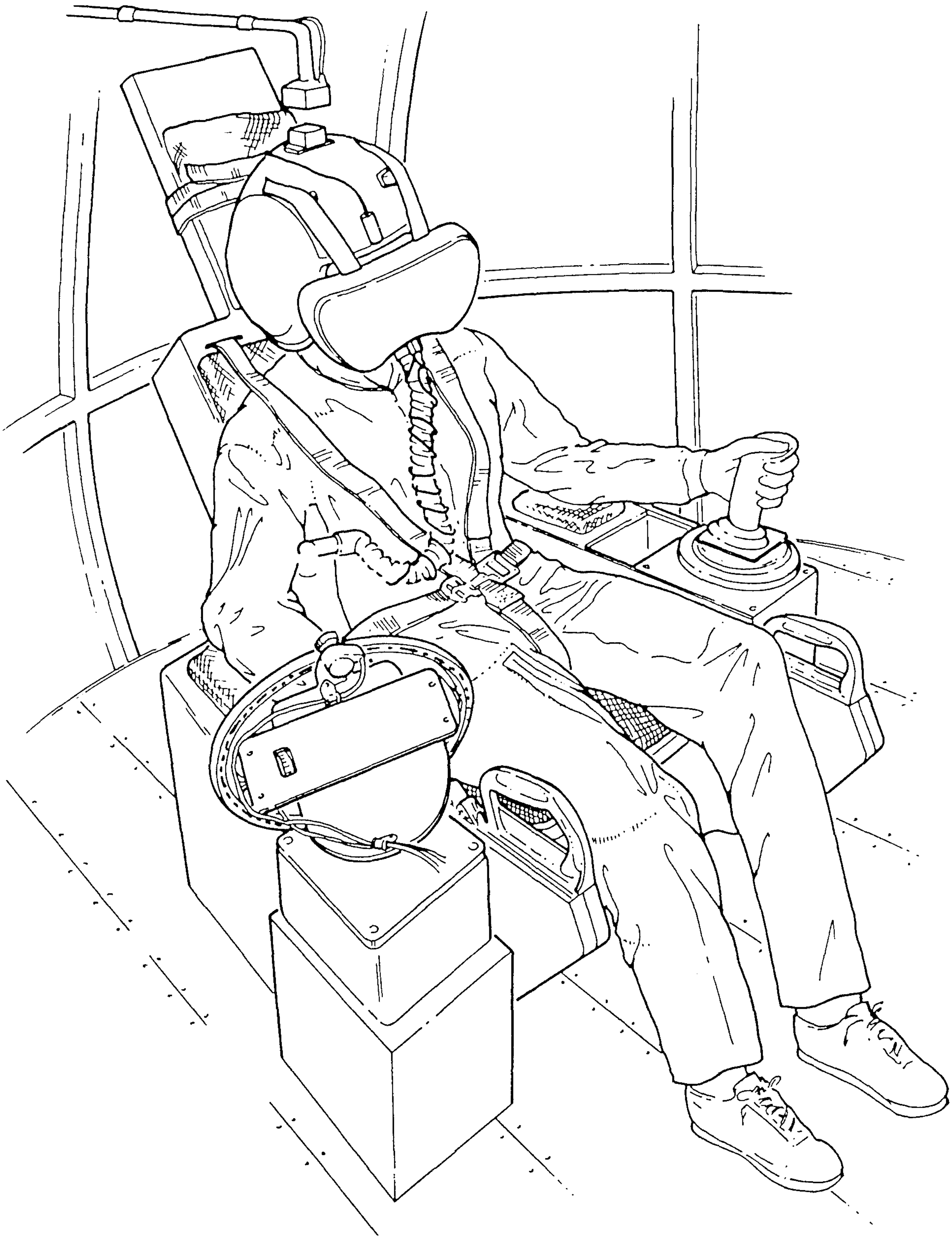


Figure 2. Virtual World of the Research Subject

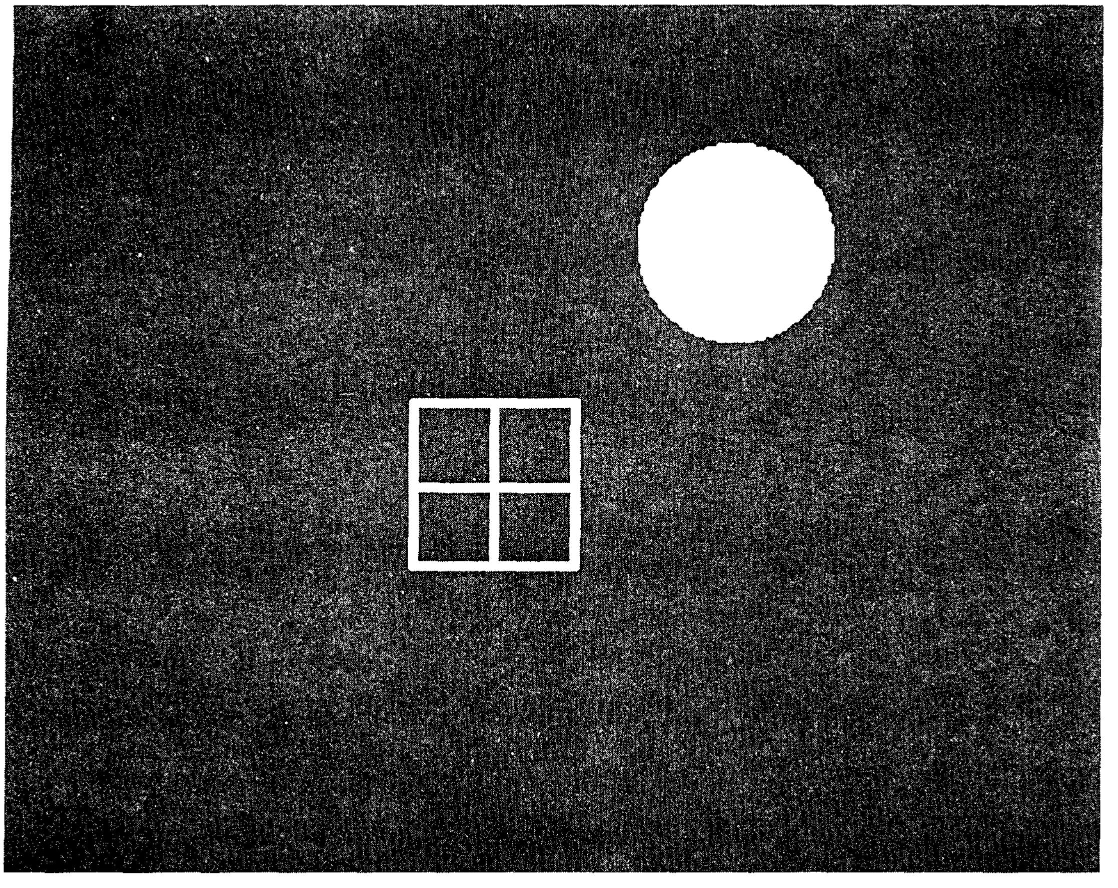


Figure 3. Task as Viewed through the Virtual Reality Goggles.

THE DRA VIRTUAL COCKPIT RESEARCH PROGRAMME

Judith Ineson
Flight Systems Department
Defence Research Agency
Farnborough, Hants, GU14 6TD
United Kingdom

SUMMARY

The aim of this paper is to describe work in progress at the Defence Research Agency (DRA) Farnborough on the Virtual Cockpit, with particular emphasis on format design and development.

The paper reviews the reasons why the concept of the Virtual Cockpit is of interest, and the ways in which it differs from the common understanding of Virtual Reality. The potential advantages and disadvantages of such a man-machine interface are discussed. The overall aims of the DRA Virtual Cockpit research programme are listed, together with a more detailed discussion of the areas of concern in the presentation of visual information.

The current status of the research programme is described. The hardware being used for this programme comprises a head-coupled binocular helmet-mounted display (HMD) system in a skeletal cockpit rig with stereoscopic, computer generated graphics, and a set of demonstration formats showing examples of the type of imagery which might be employed in a Virtual Cockpit. This is followed by a description of APHIDS (Advanced Panoramic Helmet Interface Demonstrator System) - a more capable Virtual Cockpit research rig currently being built for DRA, and of its strengths and limitations. The paper concludes with an outline of how APHIDS will be employed in the next stage of the research programme.

1 INTRODUCTION

Today's military pilot has to assimilate and interpret a vast quantity of information. As his aircraft becomes more complex and "intelligent" and the world around him becomes more dangerous, he is in danger of becoming overwhelmed with data to the point where he can no longer do his job efficiently.

The manner in which information is presented to the pilot affects his workload. The aim of a visual display is to give him the information he needs in a form which he can interpret with the minimum of cognitive effort, whilst taking his attention from his main flying task for as short a time as possible. Thus, in a fast jet, primary flight information is now displayed head-up so that the time the pilot spends looking into the cockpit is reduced. It is also collimated so that he need not spend time re-focusing his eyes. Some information, such as the pitch ladder and the bomb-fall line, which relates directly to world around him, can be seen superimposed on the real scene, facilitating the interpretation of the display. The head-up display can also relay a sensor image to the pilot, giving him a view of the scene ahead of him, which improves his night and bad-weather flying capability.

The head-up display has, however, a limited field of view and is fixed to the aircraft, so that the information is available only when the pilot is looking ahead. The next logical step was to mount the display on the helmet so that the pilot always has flight information available. Displaying a sensor image which is pointed in the same direction as the head produces a visually coupled system, such as the IHADSS system which is in service in the Apache.

A visually coupled system, consisting of a wide field of view, binocular helmet display together with good quality image generators and display sources, offers the potential to create a virtual cockpit, where most or all visual information is delivered to the pilot on the helmet display. It is hoped that the pilot's workload can be reduced by making as much information as possible spatially appropriate by projecting it so that it appears to be placed in the outside world or within the cockpit. A binaural sound system could likewise place sounds correctly in the world around the pilot. Other information which is important to the pilot can be fixed on the helmet display, so that it is always available. Pictorial data can be used instead of digital or textual if it makes the information

Presented at an AGARD Meeting on 'Virtual Interfaces: Research and Applications', October 1993.

more easily interpretable. Colour or spatial depth can be used to classify information so that time spent in searching for information can be reduced.

Because so much information has been moved onto the helmet display, ways of interacting with the information will need to be explored. There has already been a substantial amount of research in the field of direct voice input (DVI) for airborne applications, and it is likely to be an important part of the virtual cockpit. However, head and eye direction and finger position can be tracked and might be useful alternative methods of controlling the aircraft systems.

The virtual cockpit offers many potential benefits. With a visually coupled system, information can be overlaid on the real world so that features can be highlighted or cued. The field of regard is unlimited and the pilot can "see" through the aircraft structure, thus enabling the pilot to maintain visual contact with objects which would not be visible in a normal cockpit. Other non-world related information can be made static on the helmet so that it is always visible, or placed in a particular direction relative to the world or the aircraft, or kept stationary at a given point within the aircraft. This ability to lock the information to a relevant frame of reference should provide a natural, informative and interactive interface between the pilot and elements of the world around him.

There are also many potential problems to be overcome. Poor physical or optical characteristics, although tolerable in the laboratory, would make a helmet-mounted display unusable in flight by producing double images and eye fatigue. There have been some reports of misinterpretation of information and disorientation whilst using helmet-mounted displays, due to confusion between head and aircraft movements. Inadequate display resolution will result in large or illegible text and symbols. The quantity of information which will be displayed will need careful management to avoid clutter.

It is worth mentioning the ways in which the virtual cockpit deviates from the accepted concept of Virtual Reality. Virtual Reality attempts to create a compelling, synthetic, interactive world which replaces the real world. The user feels that he is part of the artificial world, and great emphasis is placed on the exclusion of the real world from the user's conscious mind. He interacts with the artificial world very much as he would the real world; by sight, sound, voice, movement and touch, but, unlike the real world, his actions are physically inconsequential.

In contrast, the aim of the virtual cockpit is to augment the real world by supplying more information than the pilot can derive from his view of the outside scene. It is a tool to help him to interpret what is going on around him, but it may be subject to error and misinterpretation. We do not wish to convince the pilot that his imagery is real or to make him feel that he is in any way insulated from the real world. His actions have real and immediate consequences both to himself and to the world around him. Virtual Reality and the virtual cockpit employ similar hardware and software techniques, and so appear superficially similar, but the fundamental aims of the two concepts are in opposition and must not be confused.

2 AIMS OF THE DRA PROGRAMME

The virtual cockpit is not a new concept, and it has been a research topic for some time. The DRA programme has so far been mainly concerned with advancing the enabling technologies sufficiently to produce equipment of high enough quality to allow an adequate exploration of the idea, and this interest will continue. However, equipment is now becoming available to start serious research in ground rigs. The first objective of this work must be to demonstrate the basic utility of the concept, that is whether or not the virtual cockpit has the potential to be a useful tool for the future pilot.

There are many areas of concern. The generation of a suitable set of formats to be presented to the pilot on the display are clearly a major part of the research programme, and there are many potential perceptual problems which could interfere with the easy use of the display. Ways of colour coding information so that it can be more easily found or so that it attracts attention is under continuing investigation, and now stereoscopy will add a new dimension to the debate. The investigations into the new control methods mentioned above will be vital to make the HMD part of an interactive

tool. 3-dimensional sound cues could be a useful aid to help the pilot to relate sounds correctly in the world around him when applied to, for example, warnings or a wingman's voice.

The programme is exploratory, and is largely aimed at providing suggestions about the viability of the concept rather than laying down rules, and one of the results of the initial DRA programme will be to suggest a variety of usable solutions to formats and control problems so that they can be investigated more thoroughly in later phases of the programme. It is intended that the programme should supply practical advice for the specification of in-service virtual cockpit systems, for example field of view, resolution, display quality, whether colour is necessary or desirable, the power of the image generator, the maximum tolerable lags in the system, the resolution and accuracy of databases and the allowable mis-match between displayed and real features.

3 THE FORMAT DEVELOPMENT PROGRAMME

My main interest is in the display formats: what information to give the pilot and how best to display it. Figure 1 shows an example of the type of image which might be used during low level flight in a fast jet, when the pilot's view of the terrain is constrained by operational or meteorological factors. Spatially and functionally, the image can be split into three major areas, as shown in Figure 2: the pilot's "head-out" view of the terrain over which he is flying, the primary flight data, and the "head-down" images containing the tactical and systems overviews. Each of these components of the overall image is intended to supply specific information, however each also raises a collection of uncertainties which must be resolved before a reasonably optimised set of display formats can be created.

3.1 Head-out scene

The head-out scene is the pilot's view of the terrain over which he is flying, and the contents of the image will depend greatly on the flying conditions. In Figure 1 there is a synthetic terrain with associated navigational and tactical features surrounding a sensor image in the centre of the picture, which provides a scene which is correct for the pilot's head direction. Also present, at the left side of the picture, is the image from a magnified, narrow field of view sensor which is locked onto a target on the ground. Areas of concern include

a) How to draw the terrain overlay and its associated features under various meteorological conditions.

b) How the synthetic features are to be merged with the sensor image, or the natural view, which features should be included and whether this selection depends on the visibility of the terrain.

c) How much mis-registration of the synthetic imagery with the real world can be tolerated, and what resolution is required for the databases used to draw the imagery.

d) Whether stereoscopy is of use in general flying, or only for specific parts of the mission, such as landing or refuelling, when the external objects are sufficiently close to have a discernable stereoscopic disparity.

d) How much image distortion is tolerable, for example, geometric differences between the images seen by the two eyes could give the user a misleading view of the world, and fusional difficulties could result in short- or long-term visual strain.

3.2 Primary Flight Data

The primary flight data replaces the information currently shown on the head-up display, and includes such items as attitude, vertical and horizontal speed, height and velocity vector. In Figure 1, attitude is given by the dots at pitch and heading intervals of 10°, and heading can be read from the compass values on the 0° pitch line. These give the effect of flying inside a sphere which is positionally centred on the aircraft but which is rotationally static in the outside world. Height and speed are shown as scales at the top of the picture, and these are fixed within the helmet display so that they are always visible.

Some of the areas we wish to examine are:

a) Which information is needed all

the time, and which should be available on request.

b) Which information to place in which frames of reference, for example attitude information should be space stabilised (stationary with respect to inertial space), but should height and speed be helmet stabilised (fixed within the helmet display), and so be always available, or aircraft stabilised (stationary within the aircraft), giving a strong cue to the forward direction.

c) How to make the different frames of reference unambiguous, so that head and aircraft movements can be differentiated.

d) Whether the use of a small stereoscopic disparity would help to separate the primary flight information from the background, and whether colour can help the pilot to select and assimilate the information.

3.3 Head-down imagery

The head-down imagery replaces the cockpit instrument panel. It will contain items such as systems monitoring and management displays, weapons management displays and plan and perspective maps, all of which are not part of the pilot's view of the external world. Below the sensor image in Figure 1 there is a perspective map, looking down on the aircraft and the surrounding terrain, with overlaid tactical and navigational features. Below this is a weapons control format and a systems summary.

The pilot will be able to call up the head-down displays at will and, assuming the use of stereoscopy, position them stably wherever he pleases within the virtual cockpit.

All of the formats need to be designed with due regard for the limited resolution of the HMD, which will affect the size and legibility of text and symbols, and on the size of the formats in the field of view. The aim will be to create easily interpretable formats which need as little visual or cognitive attention as possible.

Also of interest is the positioning of the formats in depth. Formats placed close to the pilot will be perceived as items separate from the rest of the imagery, and will also be within easy reach if a finger tracker is employed. Set against this is the time needed to alter the convergence of the eyes, and this will need investigation.

3.4 General points

There are many peculiarities of computer generated imagery which could become important when viewed on a binocular helmet, as well as some problems specific to helmet displays. Some of the main questions to be addressed include:

a) What is the effect of the raster structure of the display on the legibility of text and on the perception of stereoscopically presented imagery, and will part or all of the image will benefit from antialiasing to disguise the raster structure.

b) How the perceptual disturbances caused by the finite update rate of the display can be overcome. It is possible to perceive single objects as multiple objects, and there can be an apparent shearing or smearing effect in the image caused by head movements during the finite frame period of a scanned display device.

c) How to minimise the effects of latency, such as swimming of the image and mis-registration of the image with the outside world.

d) What is an acceptable field of view for the HMD, and if this is achieved by using optics with a partial overlap, how to minimise the perceptual problems caused by the brightness discontinuity at the boundary.

e) There are several things which could affect the visual comfort of the helmet, such as the mis-match between focus and disparity when using stereoscopic imagery, the effect of seeing a close cockpit interior behind the distant helmet imagery and whether the distant imagery should be collimated or displayed closer than infinity.

f) How can the judicious use of colour and stereoscopy help to reduce clutter and classify information, thus making the series of overlaid images more

easily interpretable.

4 THE CURRENT STATUS OF THE DRA PROGRAMME

Whilst waiting for the development of more sophisticated hardware and software, Flight Systems department has built a facility in-house in order to provide a demonstration of some aspects of the virtual cockpit and to start investigations into the formats.

4.1 Hardware

The hardware consists of a frame containing a seat, sidestick, throttle, HOTAS switches and a Head Position Sensor And Loading Mechanism (H-PSALM) which is attached to a head-mounted display. This is known as the VEIL rig (Virtual Environment Integration Laboratory) (Fig 3).

The display currently in use is the "Tin Hat", which was built in-house and uses a pair of commercial colour LCD displays as image sources. The images are binocular, with variable overlap, a brightness of 15 ft Lamberts, and the transparent combiners give a see-through of 50%. Each ocular has a 23° x 17° field of view with a resolution of 200 x 300 pixels.

Also available is a binocular helmet-mounted display built by GEC, which has a 50° field of view and better resolution and dynamic range than the Tin Hat, but which is monochrome. H-PSALM is being modified to accommodate the helmet, which is larger than the Tin Hat.

Both of the head-mounted display systems take standard 625 line, PAL video signals as inputs, one for each eye.

The computer hardware for the demonstration package consists of a Sun 3/260 workstation and two Silicon Graphics Personal Iris workstations connected by Ethernet, as shown in Figure 4. The Silicon Graphics machines are dedicated to producing the imagery for the two eyes, and run identical software with slightly displaced viewpoints to generate the appropriate disparities for imagery which is intended to be stereoscopic. The Sun runs the aircraft model, collects data from the controls and the head tracker and calculates head position and orientation. It also generates data for the head loader and sends aircraft and controls data across the Ethernet to the Silicon Graphics machines.

4.2 Imagery demonstration software

A software package was written to demonstrate the essential components of virtual cockpit imagery. The software currently runs on the VEIL rig, but the graphics software would transfer easily to any Silicon Graphics hardware. Figures 5-9 show some general pictures of the imagery, and the following sections describe the component formats in more detail. It should be noted that the figures are monochromatic reductions of colour pictures, and have lost fine detail and colour contrast.

The figures are drawn with a field of view of 48° x 36°, except for the control format in figure 9, which has a field of view of 23° x 17°.

4.2.1 Head-out scene

The outside world was limited to a purely synthetic scene generated from a rectangular grid of spot heights - the generating function is either sinusoidal or flat, although work is in progress to include a database from real terrain containing simple cultural and tactical features. The graphics machines do not allow texture to be applied easily, and so this aspect was not explored. Three types of terrain representation can be demonstrated - patchwork, height-shaded and sun-shaded (Figs 5-7).

Stereoscopy was not included in the terrain drawing for two reasons: given the coarse resolution of the display it is unlikely that any stereo effects would be visible unless the observer was very close to the terrain, and to allow the demonstration of the visual separation of the primary flight display from the background world by placing them in separate depth planes.

4.2.2 Primary flight display

Two examples of primary flight displays are included in the demonstration. The first is a simple head-up display (Fig 7), which uses a pitch ladder as the attitude indicator. The velocity vector is the aircraft symbol in the top left quadrant, and this is used as the reference point for the other symbols. Height and speed are displayed using counter-pointer dials, above the velocity vector, and between them is a 5:1 scaled heading tape. To the right and left of the

velocity vector are the vertical speed and angle of attack scales. This format is aircraft stabilised, and so appears fixed in the forward field of view as though displayed on a traditional head-up display.

The second format is based on an attitude "birdcage", giving the pilot the impression that he is flying inside a transparent sphere marked with lines of elevation and heading. Height and speed can be shown either as aircraft stabilised dials (Figs 5, 6), as in the HUD format, or as a pair of head stabilised bars showing deviation from the demanded height and speed rather than absolute height and speed (Fig 8). The bar changes colour when the deviation is greater than preset limits. The actual value of speed and height is shown alongside the bar. The aircraft symbol shows the velocity vector, and the bar on the aircraft symbol the vertical velocity.

Both formats include lines drawn parallel to the edges of the display to define the head axes, so that head and aircraft movements can be readily differentiated.

The primary flight display formats can be separated from the terrain background by adding a stereoscopic separation - four levels of disparity, and hence depth,are available within the demonstration package.

4.2.3 Head-down imagery

Due to programming complexity and the constraints of the available hardware, only two simple head-down displays are demonstrated. Both are effectively flat panels placed about a metre in front of the pilot. The first is a flight control monitoring format (Fig 5), which contains a plan of the aircraft's control surfaces with nearby pointers and scales to show their movements.

The second format (Fig 9) is a simple menu which allows the user to alter the formats in the demonstration package, for example changing the terrain type, the amount of stereoscopic disparity on the primary flight display, and switching the flight controls format off and on. The format is called up and controlled using switches on the control grips, and the functions available are shown in Figure 10.

5 THE FUTURE PROGRAMME

The facility described above demonstrates a range of aspects of the virtual cockpit, including several methods of drawing terrain, some ideas for displaying primary flight data, and for presenting monitoring and control information. The basic concept of a background terrain overlaid with 3-dimensional information is demonstrable, and is useable within the constraints of the ground rig. The effectiveness of even a small stereoscopic separation as an aid to differentiating different types of imagery can be clearly seen when using the rig.

Also demonstrable are various detrimental effects, such as the loss of textual information due to poor resolution and colour range, the restrictions of a narrow field of view and the irritations of a slow update rate. These result directly from the quality of the present hardware, and clearly much of the future work on the virtual cockpit must await better equipment.

In the meantime, the present equipment will allow useful work to continue. Experiments are planned to look at the legibility of different fonts when drawn rotated on a raster display, and on the effects of antialiasing and a stereo separation on the legibility of text. Improvements in the hardware will include a new head-mounted display with better image quality and field of view, which will not only give a more compelling demonstration of the formats, but, with the inclusion of suitable databases, will allow work to start on large scale features such as flight path marking. In order that work can continue on the primary flight display, the H-PSALM rig is being modified to take the GEC helmet, and this will also enable some comparisons to be made between colour and monochromatic formats.

The Ministry of Defence is funding the design and build of a more sophisticated virtual cockpit rig known as APHIDS (Advanced Panoramic Helmet Interface Demonstrator System), currently under construction by GEC-Marconi Avionics at Rochester. It is hoped that APHIDS will prove adequate for much of our future work: the helmet-mounted display will be a full colour, 60° field of view system, and the powerful image generation system and fibre-optic, reflective memory data

transmission system should provide vastly superior imagery to the VEIL rig, with a much reduced latency. The APHIDS hardware also includes head and eye tracking, DVI and a 3-dimensional sound system. The software includes off-line tools to help with format design and a real-time system to manage the hardware, fly the aircraft model, control the mission and draw the imagery.

The formats developed on the VEIL rig will be transferred to the APHIDS cockpit so that format development can continue with the better display equipment and the formats can be tested under a more representative environment. Work will also take place to study alternative control methods, used in conjunction with the formats.

APHIDS at present lacks any way of tracking finger position, but a finger tracker is being developed under a separate programme and this will eventually be included in the APHIDS rig so that the designation of virtual switches by hand can be compared with the other control mechanisms.

APHIDS also does not include a method of importing or synthesising a sensor insert into the computer generated scene, as illustrated in Figure 1. Since there are likely to be both technical and perceptual problems in dealing with such sensor inserts, this is seen as a deficiency, which we will attempt to rectify in the future. Also missing is a way of presenting an outside scene beyond the helmet, so that ways of merging the synthetic imagery with the real world, in different weather conditions, can be addressed. It is however hoped that this facility can also be added during future hardware development.

6 CONCLUSIONS

Flight Systems Department of the DRA has an active research programme aimed at investigating the potential benefits of the virtual cockpit. Both technological and human factors problems are being considered. A low technology research rig, VEIL, has been built, which runs a "flying" demonstration of some candidate formats, and which is now being used for more detailed work on format design. Improvements to the rig are in hand, and it is anticipated that it will remain useful as a prototyping tool.

Work is proceeding on the more advanced APHIDS hardware and software. When it becomes available, these formats can be refined and tested more rigorously. Research into control mechanisms will become an important part of the programme, and when sufficiently mature, will be combined with the formats to be tested in a mission context.

ACKNOWLEDGEMENTS

This work is being carried out with the support of the Procurement Executive of the Ministry of Defence.

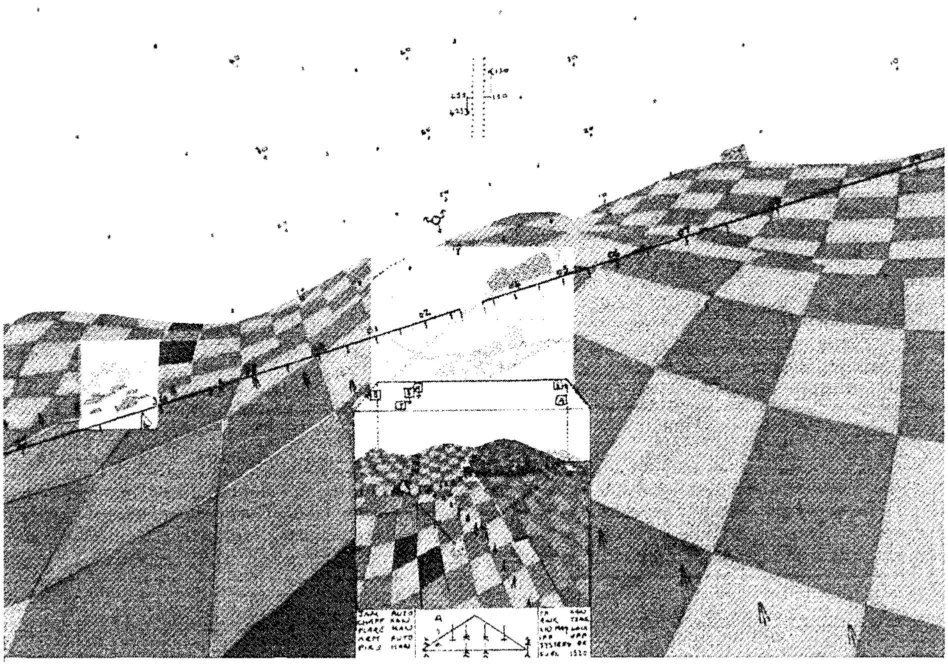


Fig 1 Example of a low level flight format

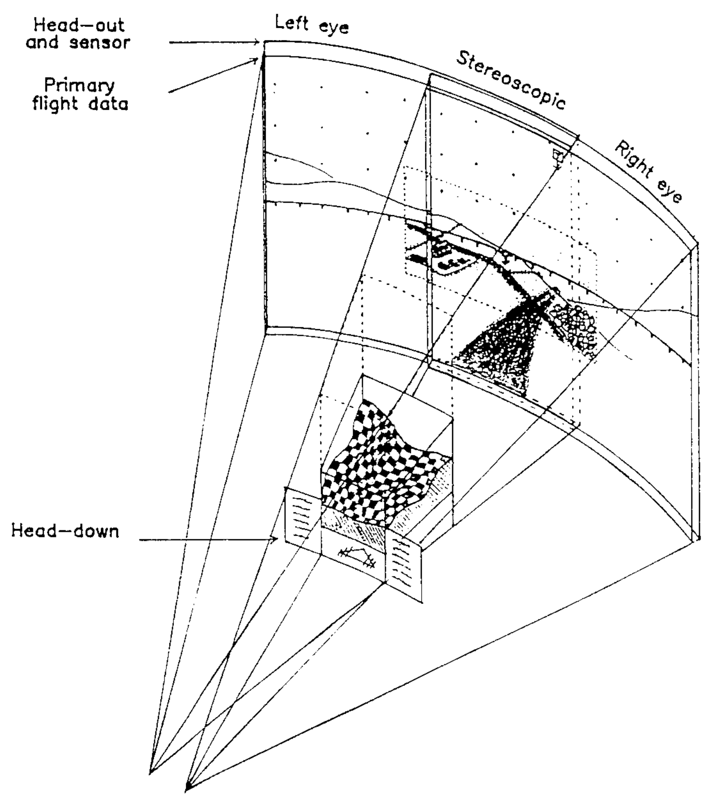


Fig 2 Spatial breakdown of the image

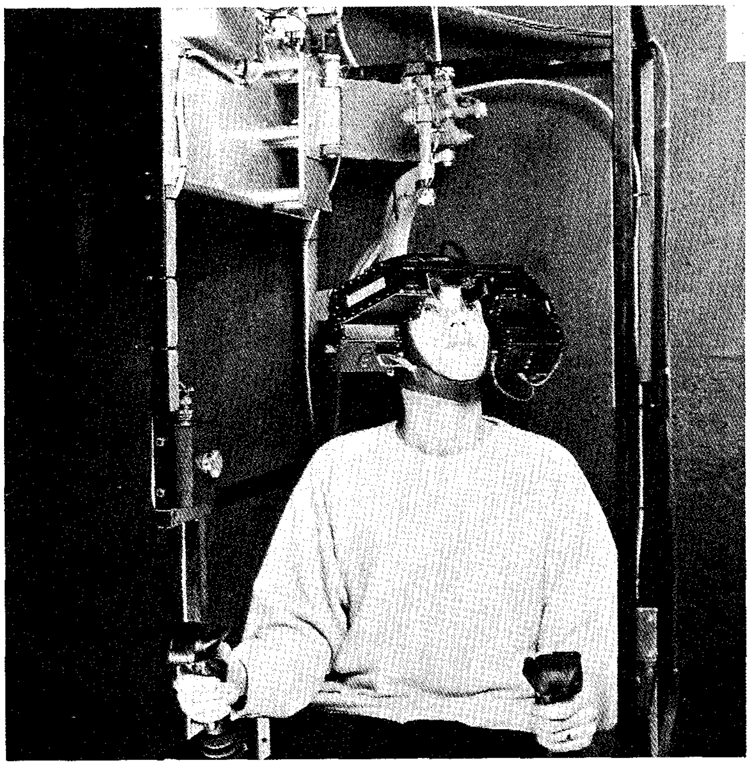


Fig 3 The VEIL rig

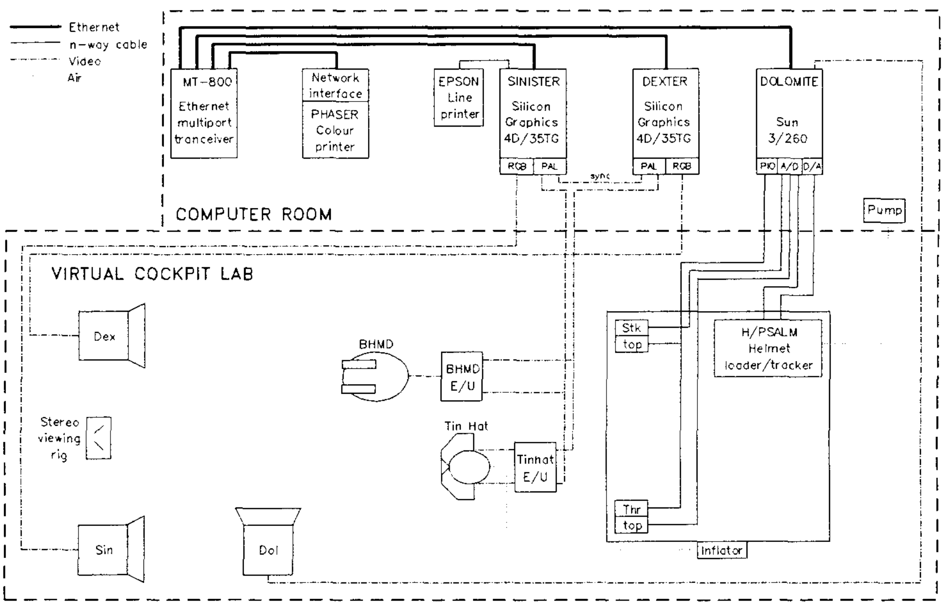


Fig 4 VEIL hardware system diagram

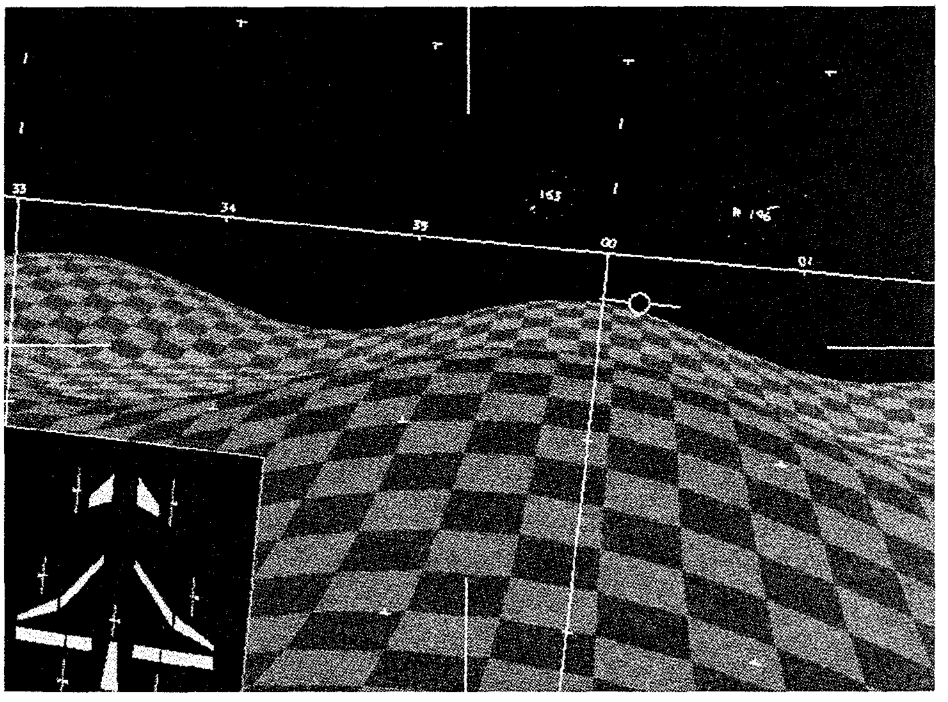


Fig 5 Patchwork terrain, birdcage primary flight format, flight controls monitor

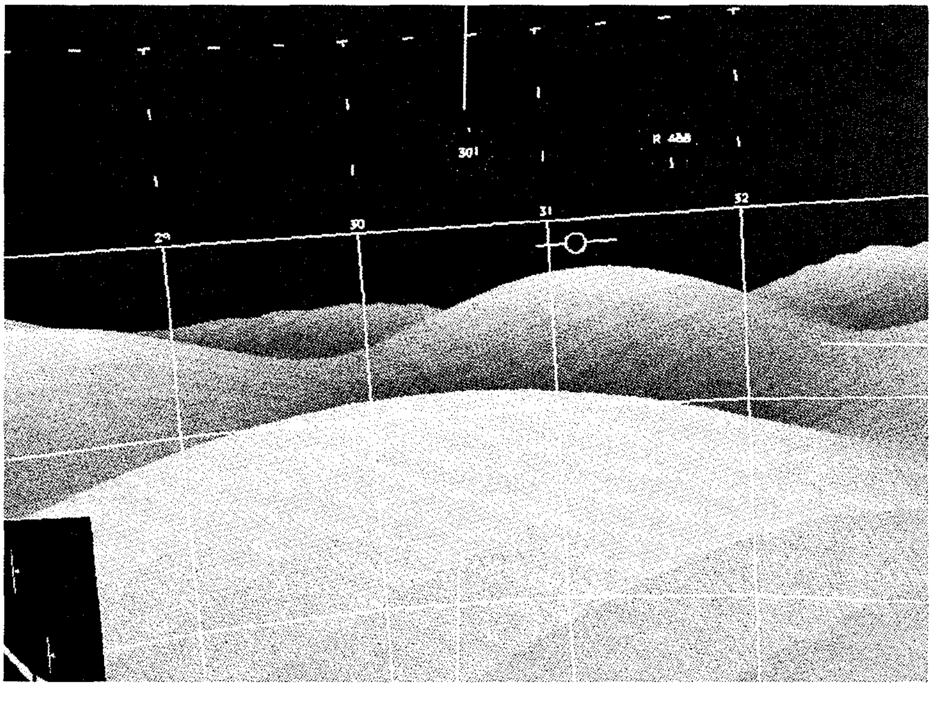


Fig 6 Height-shaded terrain, birdcage primary flight format

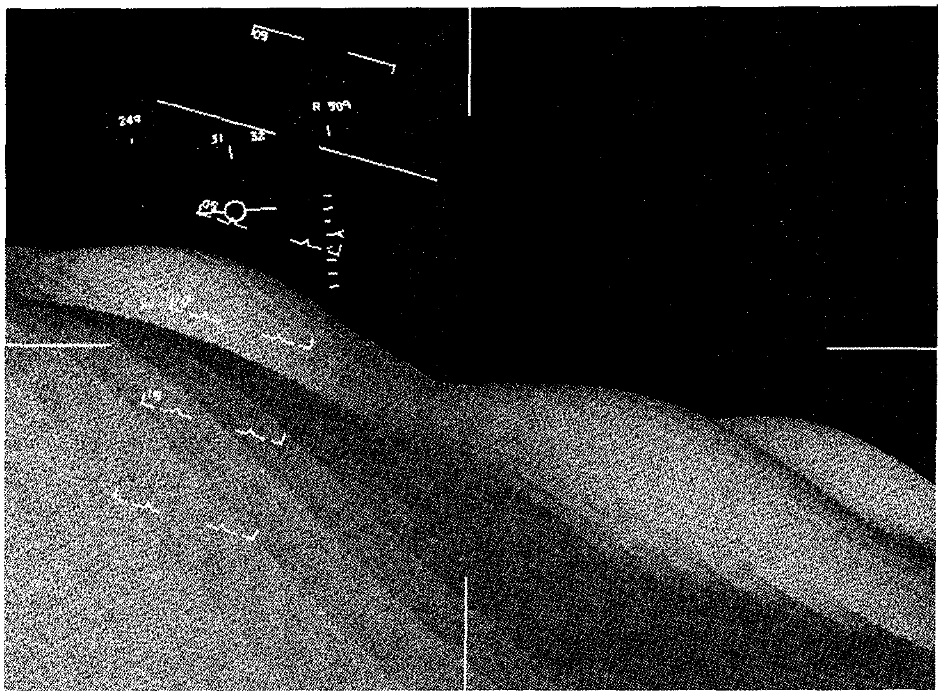


Fig 7 Sun-shaded terrain, conventional head-up display format

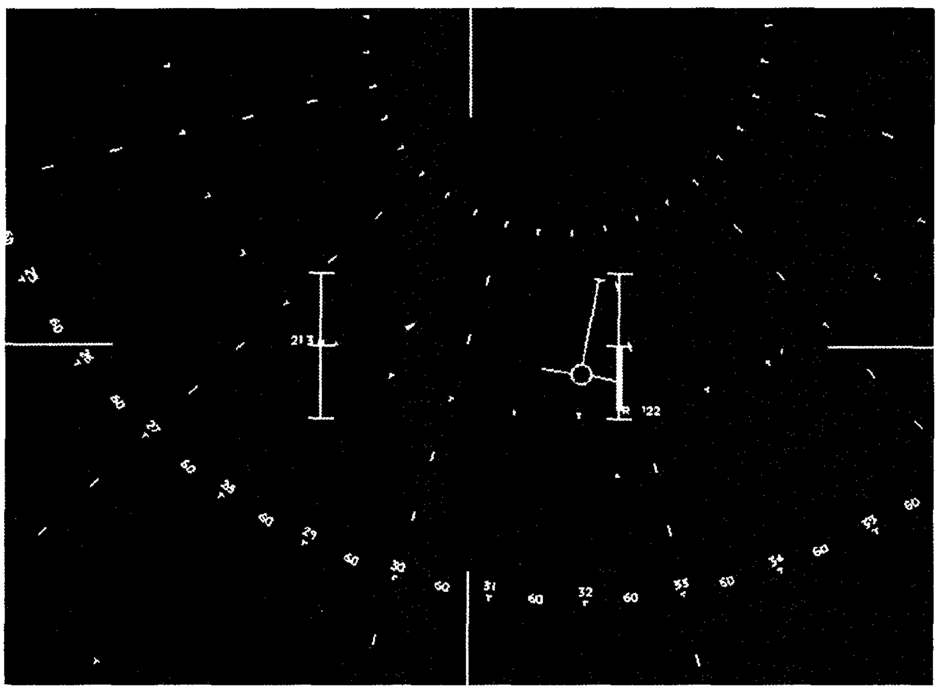


Fig 8 Points birdcage, bar-type height and speed, low height warning, aircraft 72° pitch up, climbing, looking slightly left

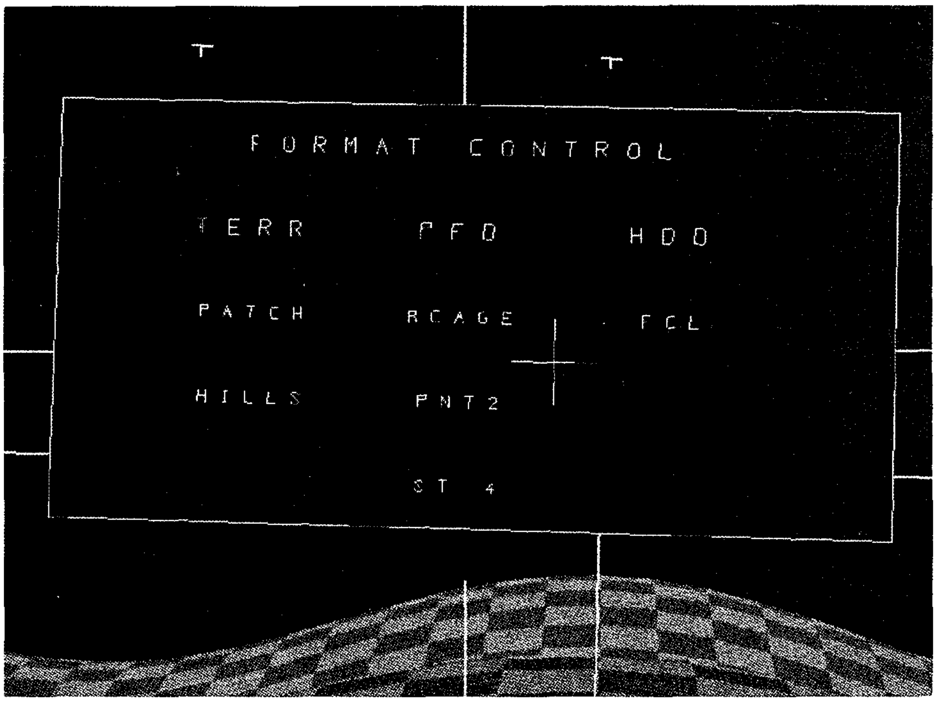


Fig 9 "Head-down" format control format

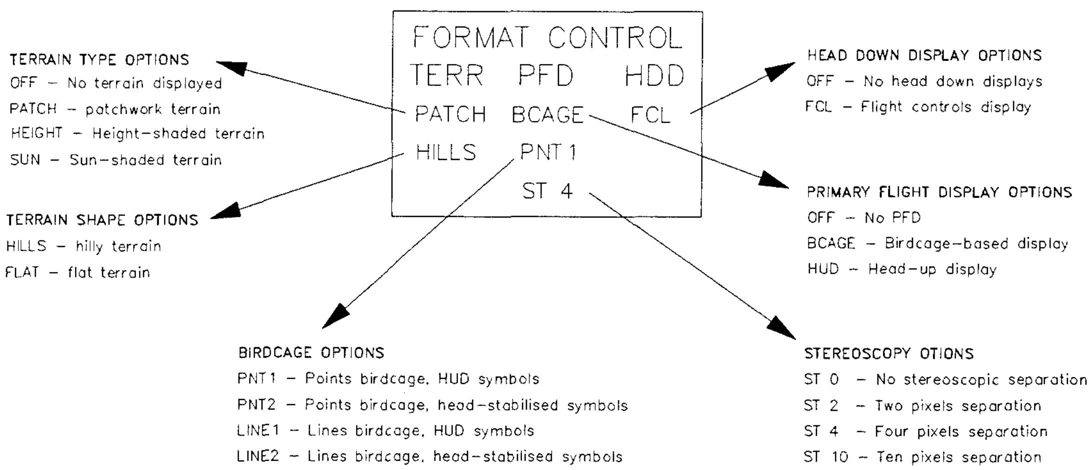


Fig 10 Format options

Virtual Reality Evolution or Revolution

Charles Grimsdale
Managing Director
DIVISION Limited
19 Apex Court
Almondsbury
Bristol
UK
BS12 4JT

There is a growing body of research which can now lead us to a strong rationale for Virtual Reality as the next generation of Human Computer Interface. As an interface metaphor Virtual Reality clearly has great potential, throughout industry, commerce, and leisure. But how will it gain acceptance. It is my belief that this will be a process of evolution rather than revolution. Much has been written about the limitations of underlying computer systems, and 3D peripherals but there is a fundamental need for more powerful and flexible software upon which to build this new generation interface.

What is fundamental about VR as an interface

I think it is valuable to first analyse the fundamental advantages of a true Virtual Reality.

3D perception - *what you see is what you get*
The shape of objects and their interrelationships remain ambiguous without true three dimensional representation. As evidenced by the art of MC Escher, perspective projection onto flat surfaces can be highly ambiguous. VR removes this ambiguity, and as a result VR represents a fundamental objective of the design process. What you conceive is what you get. Of particular importance is the sense of scale which can only be conveyed by immersing the designer in the "design". There is no longer any distinction between the object in design and reality.

Proprioperception - *the essence of experience*
The natural relationship between a movement of the user and the perceived result is critical. Simple 3D construction experiments clearly demonstrate the power of VR as a design tool. There is no longer any need for translation between the interface space and the object space, and 3D manipulation becomes trivial. So for example in mechanical construction or molecular modelling when assembling hierarchical parts in 3D, the user sees great increases in productivity from operating in the Virtual Environment.

Communication - *a shared experience*
VR promises to completely revolutionize the use of computers for co-operative working. Natural human interaction is not achievable in two dimensions. The telephone, or video phone, are effective but not absolute. Once participants share a common space they have ultimate freedom to communicate ideas.

Evolution not Revolution

How can we best exploit these clear advantages and get Virtual Reality in real use. We must_offer a path of lowest resistance - *Evolution not Revolution.*

If we take for example the Industrial design process. There is a growing demand for advanced tools to shorten the design cycle, and enable companies to bring new ideas to market. For example imagine:

1. A 3D sculpting system, which enables rapid conceptual design, e.g. shape modelling.
2. An environment modeler, which allows you to place the design object in

Presented at an AGARD Meeting on 'Virtual Interfaces: Research and Applications', October 1993.

context. e.g place the CAM shaft in the CAM guides. Place the Hi Fi in a real living room. Place the new desktop computer on a typical desk.

3. A multi user design environment, in which engineers, managers and customers, can study and discuss a design, all immersed in the same virtual environment.

Each of these capabilities could be easily achieved by providing extensions to **existing** industry standard tools. What is required is a flexible software toolkit with integrated 3D peripherals at a reasonable incremental cost, running on standard platforms. Simple tools can then be built to wrap around existing packages such as 3D Studio, Wavefront, or CATIA. If an optional VR extension to such packages were available at low cost the market demand would naturally be considerable. The technical advantages of a VR solution are clear, the commercial advantages follow naturally. However compatibility with existing market leading tools is an essential beginning.

Over the last three or four years the Virtual Reality industry has been in an early Research Phase. Many different groups have looked at what might be possible. The primary focus has been on evaluating new interface devices, such as gloves, wands, 3D mice, stereo visual systems, tactile displays etc.. and studying the underlying metaphors and psychometrics of the VR interface. Virtual Reality must now progress to a new phase of market acceptance. This requires stable platforms, and software which enables existing software vendors to trivially Virtualize their products, and encourages a new generation of software developers to establish advanced VR products.

A new phase in development

Much attention has been given to the mechanics of VR. Particularly the development of new 3D peripherals. The fundamental hardware requirements of a good VR platforms are in some cases still inadequate, but progress in the development of graphics, audio, and compute performance is very rapid. The reslut is that standard workstation platforms such as Silicon Graphics are becoming more VR capable, and specialised systems such the DIVISION ProVision and SuperVision platforms more affordable.

In order to establish this new generation of man machine interface, what is required is required above all else is a new generation of operating environment. A software environment that integrates 3D computer generated images, 2D images (stills, and full motion), 3D sound, and 3D control. In the way that X Windows, and Microsoft Windows provide a flexible development environment for 2D window based interfaces, which also enforces a standard look and feel, a new generation of software environment is required which provides the **foundation** for a true 3D look and feel. This software must be evolutionary building upon well established standard environments. We need a methodology which will co-exist with todays 2D interfaces, and add value where it is really required.

This software environment must be as flexible as possible, providing a completely application independent interface. We must look beyond the high levels tools, such as authoring, modelling, animation, or scripting tools, at the common Application Programming Interface (API) which underlies these tools. This is the software upon which applications are developed and must be equally viable as a programming interface for molecular modelling, as it is for architectural design, or flight simulation. This API must be supported by a powerful Runtime environment which ensures that interactive 3D applications run efficiently regardless of the target platform. This Runtime software must hide the must provide support for a wide range of 3D peripherals, and for multi-participant networked Virtual Realities.

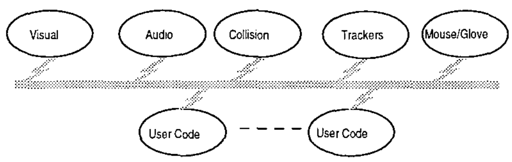
Given a well developed API and Runtime environment available on a wide range of platforms applications will start to emerge. These applications will automatically inherit a standard look and feel which will facilitate rapid acceptance within the user community.

A foundation for Progress

Over the last three years DIVISION has developed dVS, a very flexible and open software environment upon which advanced 3D interfaces can be built. dVS augments existing operating systems to provide the highest possible performance on a wide range of platforms. Based upon a distributed architecture which exploits the natural parallelism of a 3D interface, dVS is the next step in Client/Server architectures; it is a Client/Client, or Actor based architecture.

A Distributed Approach

The basic principle is to enable different components of the user interface to execute in parallel and where possible upon different processors. Defining such a distributed model greatly simplifies the process of interface development, and improves performance, regardless of whether the target machine is parallel.



Multiple dVS servers (Actors)

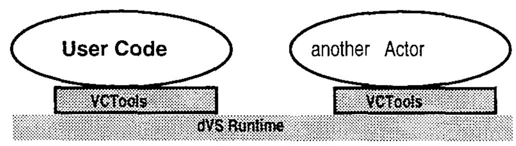
The diagram above illustrates a typical dVS configuration. The user code is quite independent of, and runs in parallel with servers (Actors) dedicated to the main display and sensor tasks of the 3D interface. dVS provides a very high level interface between user code and the standard Actors which provide visual, and audio simulation, collision detection, tracking, etc.. This level of abstraction between application software and the mechanics of the interface is very important. It ensures much greater portability of applications, and upgrade ability. Advances in graphics, audio, or i/o hardware can be exploited by enhancing only the relevant i/o Actor. The user's application code does not even need to be re-compiled. Upward compatibility is easily ensured, and performance maximized on a given platform.

The Actor model

It is very natural to decompose a complex virtual environment into a collection of completely autonomous 3D objects. dVS provides the infrastructure under which these objects co-exist. This is a very powerful metaphor, and greatly simplifies the problem of creating complex environments. The ultimate form of 3D Clip Art will become possible when we can encapsulate all attributes of an object, visual, acoustic, behavioral, etc. in a single piece of dynamically instantiable software, an Actor, which represents that object and which can be loaded at any time. Imagine an Actor which represents an autonomous automobile. This Actor responds to other objects (of known type, e.g. roads, trees, pedestrians) in a defined way, and defines the visual, acoustic, and other properties of the automobile. You then have a 3D object which can be sold to numerous customers, who want to include auto's in their virtual environments.

The API

dVS provides a very concise and powerful Application Programming Interface in the form of the VCToolkit. This is a Library of ANSI 'C' functions which manipulate high level Virtual Environment *Objects*. An *Object* contains a number of basic attributes such as *visual*, *audio*, *constraint*, and *collision* attributes. The VCToolkit also has a powerful event processing mechanism which allows call backs (actions) to be assigned to particular events. This interface is ultimately flexible and completely abstracts the developer from the underlying hardware.



Multiple Participants

Another careful consideration for the whole dVS design has been support for multiple user's, in a common shared environment. The distributed nature of dVS naturally supports this concept. The whole core of the dVS software is concerned with maintaining the consistency of data among multiple Actors, and these Actors can of course represent different users.

High Level Tools

dVS is a powerful platform-independent foundation for developing and running VR applications. However developing new applications does require programming effort. There is a clear need for higher level

tools which allow non-programmers to develop virtual worlds.

DIVISION has developed a unique Virtual World simulation and authoring package called **AMAZE**, which allows users to quickly build and experience virtual environments. This Actor runs on top of the basic dVS Runtime. Whole environments can be constructed without writing a single line of code. The base geometry of these worlds can be imported from standard CAD packages such as AutoCAD, or 3D Studio. Further attributes such as sound, animation, etc.. can then be added. This software has proved incredibly flexible, and has been used to prototype many applications from golf course construction, to molecular modelling.

Other high level tools are needed to facilitate the entry to Virtual Reality, tools for Acoustic modelling, and geometric modelling, and complex behaviour modelling. With the stable foundations of a software environment such as dVS such tools can now be rapidly developed.

Conclusions

Rapid progress within the Virtual Reality market now depends upon widespread application development and acceptance. This is most likely to be a more evolutionary than revolutionary process, with existing software vendors slowly virtualizing their products. However before existing software vendors will consider making the necessary investment they must have stable system software upon which to build. The development of dVS and other solid VR development toolkits is essential to this process of evolution.

MANUAL TRACKING PERFORMANCE USING A VIRTUAL HAND CONTROLLER: A COMPARISON STUDY

by

Robert G. Eggleston
Human Engineering Division
Armstrong Laboratory
Wright-Patterson AFB, Ohio 45433-7022
United States

William P. Janson
and
Sridhar Adapalli
Logicon Technical Services
PO Box 317258
Dayton, Ohio 45430
United States

ABSTRACT

This study compares a virtual hand controller (magnetic sensor attached to a glove) with a physical displacement stick in a single-axis manual control task. Three different control/display (C/D) ratios were used with each controller. Control performance was found to vary significantly with C/D ratio. When across-device comparisons were made at identical C/D ratios, a slight but significant performance advantage was found for the displacement stick at one C/D level. When between-device comparisons were made on the basis of a performance matching technique, the results were comparable for the virtual and physical hand controllers. The issue of how to best match test conditions to achieve an unbiased comparison of control devices is addressed. Arguments are advanced in favor of using the performance based matching technique. From this perspective, the data are interpreted to support the claim that comparable manual control performance can be achieved with a virtual hand controller.

INTRODUCTION

New technology makes it possible to produce a "virtual" hand controller. Any technique that can sense the movement of a body segment (e.g. hand, arm, etc.) and convert this data to orientation and position information at real-time rates may be used to create a virtual controller. A popular method is to use a magnetic sensing system coupled to a computer and display device to provide a closed loop control system. If the sensor is affixed to the hand, say by a glove, then a virtual hand controller is formed.

There are two general domains where a virtual hand controller may have value. First, a virtual hand controller may be a suitable substitute for a physical hand controller that is designed as an interface to a physical vehicle or system. A virtual control may provide a more natural mapping between user movements and desired changes in vehicle/system state. For example, coordinated vehicle translations and rotations could be controlled by translational and rotational hand-arm movements, respectively. This might lead to improved pilot performance with a force vectoring aircraft like the X-31 and it could lend itself to a natural mapping for translational and rotational control of a helicopter.

The second domain is the rapidly growing area of virtual environments. Here, a virtual hand controller could be used to control a virtual object, instrument, machine, or vehicle from *inside* the environment. That is, the human perceives him or her self to be behaving inside the synthetic environment. In this application, hand control could come in the form of direct manipulation of virtual objects by natural body movements, such as a set of body gestures interpreted as commands for guidance and navigation of self-movement, or as the manual operation of a virtual control device, like a stick or mouse, attached to a virtual system/vehicle. In the latter case, the distinction between these two domains (interface to physical systems and interface to virtual systems) should disappear to the extent that sensory and information feedback from a virtual device corresponds to the physical one.

Whether one is controlling a physical system, a virtual system, or self-movement with virtual methods, it is important to determine the level of performance that can be achieved by these new methods. One probably would not substitute a virtual hand controller for a physical controller unless control performance was at least as good, if not better, than with a physical method. Performance in virtual space often will not be an end in itself. The virtual environment will be used for training new skills, practicing procedures to maintain proficiency, or it will be connected through a suitable computer based system to physical instruments and systems that respond in real-time to computer interpretations of a person's movements as input commands. Thus, human manual performance in virtual space must be adequate to support effective transfer of training as well as skillful control of robotic end effectors. What level of control performance is needed? How do human performance design requirements interact with hardware, software and computational modeling properties of virtual systems?

This study is the first in a set of experiments that have been formulated to provide answers to these questions. With respect to a virtual hand controller, the first question to ask is: can a person achieve control performance comparable to that obtainable with a physical control device? In this paper, we report the results of an experiment which addressed this question by comparing control performance between a virtual hand controller, fashioned from a magnetic tracker, and a

Presented at an AGARD Meeting on 'Virtual Interfaces: Research and Applications', October 1993.

conventional side-mounted displacement stick. Control performance with these types of devices was assessed with a compensatory tracking task.

An Unbiased Comparison

Conceptually, our goal is to determine the relationship between the optimal performance obtainable with a physical controller vs. the optimal performance achievable with a virtual controller for a given control system. In order to construct a fair (unbiased) comparison, the challenge is to be able to specify control-display conditions that will produce "optimal" performance for each device. These conditions may not be identical for each actuator.

It is known that manual control performance is affected by the ratio of control element movement to the movement of the controlled object portrayed on a feedback display. A high control/display (C/D) ratio improves performance for some task conditions while a low C/D ratio is best for others. Many complex tasks require a trade-off between C/D ratios to achieve optimal performance (Ref. 1,2). In these instances, control performance varies as an inverted-U shape function with C/D ratio. Further, the optimal C/D ratio may change with properties of the control actuator (Ref. 3). We can use the inverted-U functional relationship as the basis for objectively establishing an unbiased comparison of control devices. However, the problem of selecting the proper levels of C/D ratio for each device becomes an issue. In other words, we must now solve the problem of how to match control devices in terms of C/D ratio. Since C/D ratio may interact with actuator-specific properties, this presents a new problem.

Several methods for matching devices suggest themselves: physical identity, performance based, or effort based. The Physical Identity method is straightforward; just select one or more C/D ratios and assume performance co-varies equally across device type with this variable. If the assumption is invalid, however, the resulting comparisons will produce a biased assessment. The Performance Based method requires control performance data be collected at several C/D ratios for each device type. Then, across-device matches are formed on a high-high, low-low, etc. performance basis. A preliminary experiment, or prior data, will be needed to aid in the selection of C/D levels in order to ensure they bracket the C/D range where optimal performance is expected to occur. To establish an Effort Based method, the same procedures used to define appropriate C/D ratios on the basis of performance data must again be employed, only this time an effort metric would be used. While it is not necessary for the high-high match between devices to be at the "optimal" performance point for an unbiased comparison, if the optimal point can be identified, the resulting comparison would provide absolute performance as well as relative performance information.

In the present study, the physical control device was an ordinary side-mounted displacement stick and the control task was single axis. Hence, rotation angle in the vertical plane was the controlled variable (i.e., roll axis for an aircraft). Our virtual hand controller also responded to angular rotation. Thus on the surface at least, the same C/D ratio might be expected to have the same impact on performance with both types of controllers. But under closer inspection, this expectation becomes more difficult to defend. The displacement stick is a spring loaded, return-to-center type with the point of rotation located several inches below the region where it is held by the user. The kinematic movement of the arm required by this arrangement is a rotation about the shoulder and a translation about the elbow to produce a rotation of the stick. This is not the same movement required by the virtual device. It requires a rotation of the hand about the forearm to produce the controller output, a related but different kinematic action. The two devices also differed in the magnitude of force required to produce a unit rotation output, and to recenter the controller. As a result, it is not at all clear that optimal performance for both devices will be obtained at the same C/D ratio.

Preliminary performance data has been collected from three subjects using the controllers for this study. The data clearly showed the effect of different C/D ratios on control performance, and also suggested that optimal performance would be achieved at a higher C/D ratio for the virtual device relative to the physical one, but the difference in optimal C/D ratio was not large. Thus, while the pilot data suggests a Performance Based method is likely to be a good way to match devices, since this was a small data set and differences were not large, a straight physical matching technique cannot be ruled out as an appropriate method. Accordingly, we designed the experiment in a manner that would allow us to use both methods. In addition, we collected subjective workload data that could also be used to establish an Effort Based matching procedure. However, C/D ratios were not selected on this basis and, therefore, we cannot be confident the C/D levels used actually bracket "optimal" (least) effort for each control device type.

METHOD

Experimental Design

A two-factor, within-subjects design was employed. The factors were: Device Type (physical displacement stick and a virtual device consisting of a magnetic tracker attached to a glove) and C/D ratio (0.7, 1.4, 2.1, and 2.8). These values reflect the size of rotation angle (degrees) required to translate a cursor 1.0 cm. laterally. Based on the pilot data, C/D ratios were nested under Device Type. For the physical device (STICK), C/Ds 0.7, 1.4 and 2.1 were used. For the virtual device (GLOVE), C/Ds 1.4, 2.1, and 2.8 were used. Control performance was expected to produce an inverted-U shaped curve as a function of C/D for each device type. Thus, using a Performance Based matching method, STICK 1.4 data should be compared with GLOVE 2.1 data. If a Physical Based matching method is used, then STICK 1.4 and STICK 2.1 should be compared with GLOVE 1.4 and GLOVE 2.1, respectively. Presentation order of the treatment conditions were organized according to a Latin Square to counterbalance treatment order across subjects.

Apparatus

The study was conducted in the Virtual Environment Interface Laboratory (VEIL) at the Armstrong Laboratory, located at Wright-Patterson AFB, Ohio.

Hand Controllers

Two types of hand controllers were used for this study: a virtual hand controller and a displacement joystick. The *Virtual Hand Controller* was produced by mating a standard issue Nomex flyers glove (summer type GS/FRP-2/Mil-G-8118) with an Ascension Technology Corporation "Bird" magnetic orientation/ position tracking system, model #6BI001. The sensor was affixed to the back of the flight glove, approximately in the region aligned with the center of the palm of the subject's hand. The Bird transmitter was located approximately 19 cm (7 inches) in front of the sensor. Only x-axis orientation information was processed for the single-axis tracking task used in this study. The Bird was configured in "point mode" and standard filter settings were used. Orientation information was passed over an RS-232 link to a 386 personal computer which housed the plant dynamic and image generation algorithms. From there, the error signal was sent to a black and white television monitor for visual feedback to the subject.

The displacement stick was produced by Measurement Systems Corporation, model #12494. This is a spring-loaded, return-to-center isotonic stick. Only x-axis deflections were recorded for this experiment. Analog signals from the STICK were passed through an analog-to-digital converter to the same plant, image generation algorithms, and display used with the *Virtual Hand Controller* (GLOVE).

Due to processing requirements associated with the Bird magnetic tracking system, RS-232 communications, and graphics processing, there was a transport delay of approximately 58 msec when the GLOVE was included in the system. To ensure comparability across control systems, a delay was added to the displacement stick (STICK) to equate transport delay across devices. Conversely, a nonlinearity in the STICK movement response around the center position (i.e. a dead band) had to be added to the GLOVE to match the two systems.

Cockpit

Both hand controllers were installed in a single-seat cockpit simulator. The STICK was mounted on the right console approximately 48 cm in front of the seat back. Subjects were instructed to hold their hand (GLOVE controller) in the same location in the virtual device conditions. In these conditions, the displacement stick was removed from the cockpit. An adjustable height armrest was used to provide arm support. Visual feedback for the tracking task was provided by a monochrome television monitor (approximately 19 cm by 14 cm) which was located at eye level approximately 66 cm straight ahead of the subject.

Subjects

Six right-handed male members of a contractor maintained subject pool (age range from 19 to 23 years old) served as paid subjects. All subjects were screened to ensure they did not have any visual or physical anomalies that would restrict their ability to perform the task. As an additional incentive to maintain motivation for high performance across all treatment conditions, a cash bonus was awarded to the two highest scores in each condition.

Task

The Critical-instability Tracking Task (CTT), introduced by Jex and his colleagues in 1966, was used for this experiment. This task has been used for many years as an aid to the design of manual control systems for military aircraft. It is a first-order compensatory tracking task with an unstable element lambda whose rate of change varies nonlinearly with time. The operator attempts to minimize error that is induced by his/her own actions as expressed through the unstable pole lambda. Tracking continues until a preset error magnitude is exceeded. Using the standard conditions described by Jex (Ref. 4), a typical trial lasts on the order of 20-40 seconds. Details of the plant dynamics task can be found in Ref. 5.

Because of its unstable nature, the CTT presents a challenging manual control problem. It is easy to learn and subjects with no experience with flight control systems can achieve performance levels similar to those of skilled pilots. Thus, it is believed to tap fundamental aspects of human manual control performance (Ref. 6).

It is well known that a person behaves as an adaptive controller and can compensate in different ways to disturbances and control dynamics (Ref. 7,8). Thus, strategy differences across subjects can introduce interpretation problems when assessing operator control performance. The CTT minimizes this problem by forcing the operator's dynamic behavior to a limit by making phase and gain margins progressively more stringent until control is lost (Ref. 5). Control behavior, therefore, is made comparable at this limit point, which is indexed by the magnitude of lambda. It has been shown that CTT performance is affected by the type of control device used. Performance with a force stick is significantly better than with a displacement stick (Ref. 3). Thus, the CTT should be able to detect performance differences due to device type. It is on this basis that we selected the CTT as a good task to use to compare performance between a virtual and a physical hand controller.

The CTT was implemented on a 386 IBM compatible computer. The error signal from a control device (either the STICK or GLOVE) was conditioned and fed into the plant dynamics. The generated error signal was used to drive the horizontal location of a cursor presented on the monitor. The cursor was in the shape of a + sign with arm lengths of 7.3 mm (vertical) and 8.7 mm (horizontal) and contained a 3.2 mm by 4.8 mm hole cut from its center. A reference mark, a + sign the size of the hole, was stationary at the center of the screen. Thus, when there was no tracking error, the cursor symbol and reference symbol fused

to form an unbroken + sign. A thin white circle, centered on the reference mark and 12.7 cm (5 inches) in diameter, defined the permissible error limit. Once the cursor crossed this line, the screen would go blank and, after delay of about a second, a numeric score indicating the lambda level when control was lost was posted on the screen until the next trial began.

Procedure

Subjects were run one at a time in sessions that lasted about 45 minutes each. No more than one session per subject was administered per day. A session contained 50 trials which lasted (nominally) 30 seconds each. The first 10 trials in each session were treated as warm up trials and were not included in the data analysis. Trials were delivered when the subject was ready, usually at 2-5 second spacing. A two-minute rest break was scheduled after every ten trials and could be requested at any other time. Subjects were randomly assigned to a treatment order sequence. Testing in a treatment condition stopped once tracking performance showed evidence of leveling off at an asymptote, which was identified by a preset performance criteria. The performance criterion was defined by two tests. The median lambda scores from three consecutive sessions were submitted to a regression analysis. If the first and third session's predicted values were found to be within 5% of each other, the first test was passed. The second test measured the difference between the actual median values for sessions two and three. These values had to be within 5% of each other to pass this test. The first test indicates when performance has leveled off (zero slope) for three sessions. The second test guards against nonlinearities that could escape detection by the first test. Together, they provide a stringent criteria for performance consistency. The competitive and monetary incentives (explained below) helped to ensure that this consistency is first evidenced at the highest obtainable level of performance. Test trials were administered in this manner until all treatment conditions were completed. Subjects did not know the criteria factors used to define completion of a treatment condition.

Subjects were instructed to strive for maximum performance by always attempting to minimize tracking error. While the challenge of the task frequently provides sufficient motivation by itself, two additional methods were used to promote sustained high motivation. First, high scores from each condition were posted for public inspection (coded by subject number). Second, a financial bonus was awarded at the end of the study to the two top scores in each treatment condition.

For the GLOVE conditions, the subjects were instructed at the beginning of each session to position their right hand over a fiducial mark located on the right cockpit console at the centerline mounting location for the STICK. While hand posture was not strictly controlled, subjects were asked to place their hand in a manner simulating a grip on a stick and to find a comfortable position near a vertical alignment. This hand position was entered into the computer as the zero reference point for hand-arm rotation. At the beginning of each trial. the subject had to return the GLOVE controller to this position. This alignment task was aided by the use of a small target circle shown on the display and a small + sign which showed hand position. Once this position was held for 2 seconds, the tracking task appeared on the display and hand movements were coupled to the plant dynamics.

At the conclusion of each treatment condition, subjects completed a questionnaire eliciting observations and comments about the task and their performance. In addition, the NASA Task Load Index (TLX) was administered to assess workload at the conclusion of the first and last session of each treatment condition. Workload is defined by six subscales: mental demand, physical demand, temporal demand, own performance, effort, and frustration. After a condition is rated in terms of these scales, a paired-comparison procedure using the six scales is completed to establish weighting factors that reflect individual models of workload. A detailed discussion of this workload measurement instrument can be found in (Ref. 9).

RESULTS

The conventional tracking performance measure for the CTT is the value of lambda at the time when control is lost. All analyses were performed on mean lambda score per session (40 data points) for each of the three sessions (per condition) when the subject passed the performance criterion. All subjects required six or more sessions to meet this criterion in their first treatment condition, but showed savings in subsequent conditions.

An ANOVA was accomplished on the performance data using Subject, Order, and Condition as the variables. Condition had six levels, each defining a different Device Type (virtual vs. physical) and C/D ratio combination. With this statistical design, as opposed to the nested experimental design, the expected individual differences between subjects and any residual order effect can be separated out from the main variable of interest (Condition). A significant effect was found for Subject ($\underline{F}$ (5,20) = 50.60, $p < .0001$), Order ($\underline{F}$ (5,20) = 16.51, $p < .0001$), and Condition ($\underline{F}$ (5,20) = 18.11, $p < .0001$). These three variables accounted for 95.5% of the total variance.

The slight order effect can be seen in Figure 1. A post hoc analysis of the Order variable using the Tukey HSD test (alpha = .05, critical range = 4.445) indicated that performance on each subject's first treatment was significantly less than on those in order positions 3-6. In addition, performance on the treatment administered second was significantly less than those for order positions 4-6. Thus, in spite of the conservative criterion used to estimate the cessation of training effects, some additional learning was still evident in the data.

The main interest of the experiment involves the Condition variable. Each level of this variable combined a C/D ratio and controller type. Mean performance for each condition is shown in Figure 2. Highest performance was achieved with the STICK-C/D 2.1 and GLOVE-C/D 2.8 conditions, with the difference between the two being negligible. Slightly lower but comparable performance was achieved with the STICK-C/D 1.4 and GLOVE-C/D 2.1 arrangement. The lowest scores were

obtained with STICK-C/D 0.7 and GLOVE-C/D 1.4, with the GLOVE performance being slightly higher. Based on individual post hoc comparisons using Tukey's HSD test, the following statistically significant performance differences were found: performance with STICK-C/D 2.1 and GLOVE-C/D 2.8 were both better than with STICK-C/D 0.7 and GLOVE-C/D 1.4. Also, performance with GLOVE-C/D 2.1 was better than with STICK-C/D 0.7(alpha = 0.05, critical range = 4.445 for all comparisons). All other pairwise contrasts were not significant.

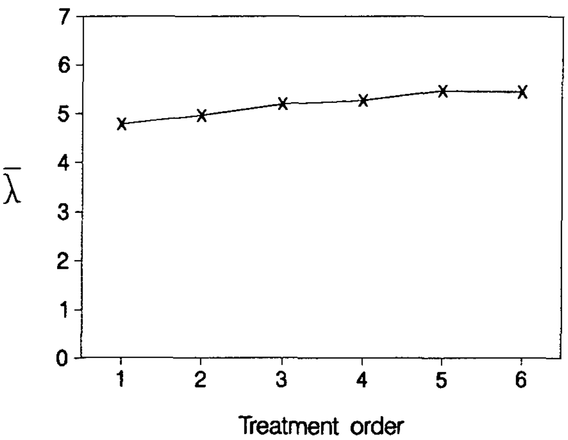


Figure 1. Mean Lambda as a Function of Treatment Order

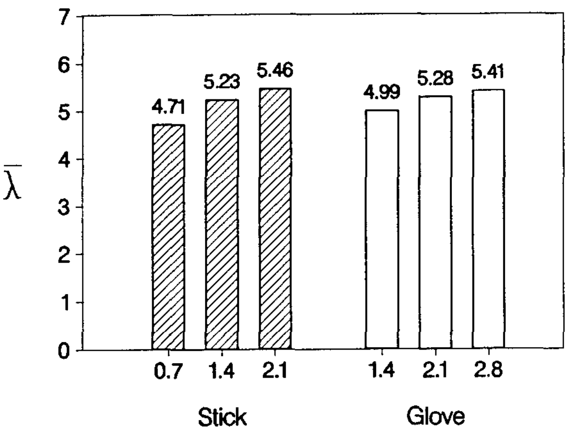


Figure 2. Histogram Showing Mean Lambda by Control Device and C/D Ratio

TLX workload ratings were collected at the end of each treatment condition. Because some subjects tended to always rate workload higher than others, these scores were normalized with respect to the STICK-C/D 0.7 condition. Normalized scores were formed by dividing each subject's scores by the value obtained in STICK-C/D 0.7 condition. Mean TLX rating scores are shown by condition in Figure 3.

Mean TLX ratings based on this normalized data were used in a 6 by 6 ANOVA (Subjects by Condition) to assess perceived workload differences across the six device-C/D ratio conditions. The Subject variable was significant ($\underline{F}$ (5, 25) = 7.01, $p < .0003$), but the Condition variable was not ($\underline{F}$ (5,25) = 2.2, $p < .086$).

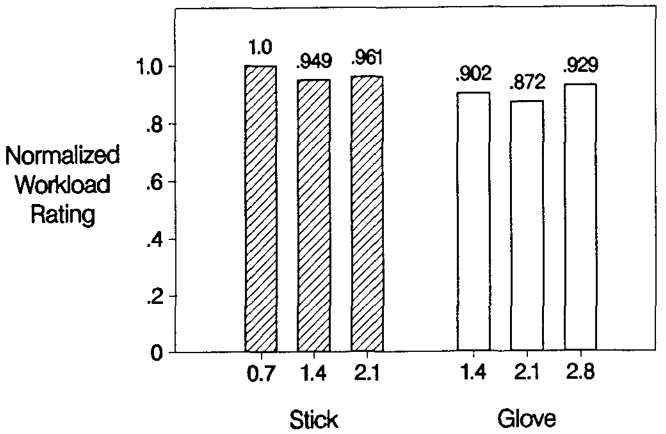


Figure 3. Histogram Showing Mean Normalized TLX Workload Ratings by Control Device and C/D Ratio

DISCUSSION

The purpose of this study was to compare manual control performance between the use of a virtual (GLOVE) hand controller and a conventional displacement stick. While the general results presented above show some between-device differences in performance, proper matching procedures must be followed before the data can be meaningfully interpreted. To aid in this process, mean lambda performance has been plotted by C/D level separately for each device, which is shown in Figure 4a. We had expected maximum performance to occur in the STICK-C/D 1.4 and GLOVE-C/D 2.1 conditions. The data in the figure show, however, that performance in conditions STICK-C/D 2.1 and GLOVE-C/D 2.8 was actually the best. Using a Performance Based matching method, therefore, these are the conditions that should be compared. However, a match on this basis may still not produce an unbiased comparison. We expected the data to show an inverted-U shape function of C/D level, with maximum performance at the middle of three conditions for each device. As just indicated, the actual maximum performance occurred at the highest C/D ratio for each device. While the data gives the hint of an inverted-U function, it does not contain the peak, which is where the between-device match should occur. More importantly, it is not clear from Figure 4a that the distance to the predicted peak GLOVE performance is in the same relative relationship as that for STICK performance. As a result, even if we match on high performance across GLOVE and STICK, the comparison may still be biased.

To address this problem, the data points for each device were fitted to a quadratic equation, as shown in Figures 4b and 4c. Based on the quadratic function, "optimal" performance is predicted to occur at a C/D level of approximately 3.1 for the GLOVE and 2.3 for the STICK. Since the relative distance between these values and the maximum performance points for each device is about the same (0.3 lambda for the GLOVE and 0.2 lambda for the STICK), these conditions may be regarded as approximately matched. Thus, any measurable performance

differences between these conditions can be regarded as due to the unique properties of the hand controllers alone (plus measurement error), since all other factors were held constant.

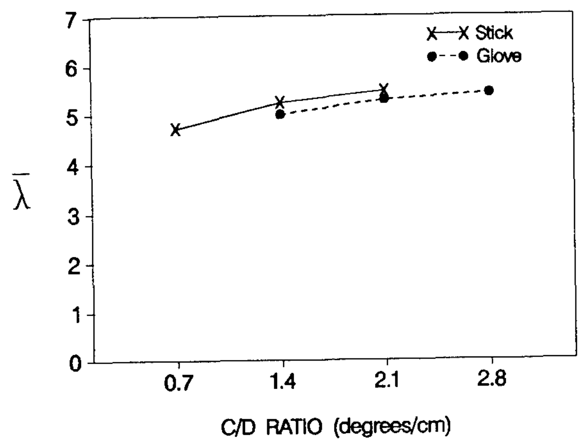


Figure 4a. Mean Lambda by C/D Ratio, Nested Within Control Device (GLOVE and STICK)

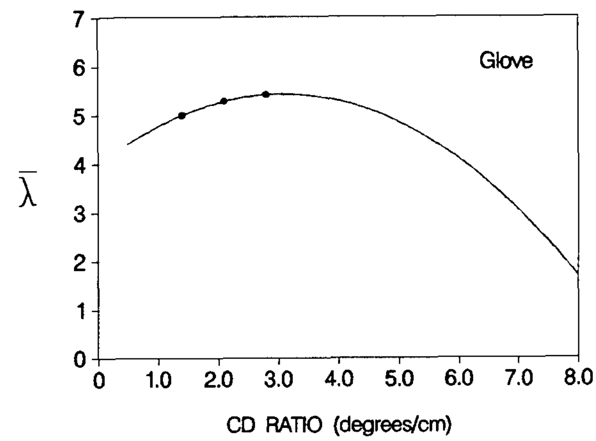


Figure 4b. Quadratic Curve Fit to Mean Lambda Data for GLOVE Control Device

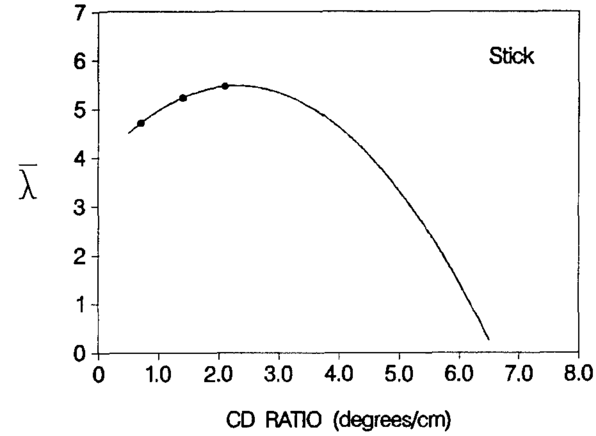


Figure 4c. Quadratic Curve Fit to Mean Lambda Data for STICK Control Device

As the post hoc comparisons showed (see Figure 2), this difference was not significant. Therefore, according to a Performance Based matching method, the results argue that it is possible to obtain comparable control performance with a virtual hand controller.

It should be noted that there is risk in extending this analysis to the mid and low performance data. A visual inspection of Figures 4b and 4c indicates that distance below the peak of the curve for the mid and low STICK conditions is greater than that for the respective GLOVE conditions. Thus, across-controller device matches at these points may be biased.

Based on a Physical Identity Based matching method, GLOVE-C/D 2.1 / STICK-C/D 2.1, and GLOVE-C/D 1.4 / STICK-C/D 1.4 conditions represent valid comparisons. These comparisons are shown in Figure 5. There is a slight mean performance edge in both comparisons for the STICK, which is statistically significant for the C/D 2.1 comparison. Thus from this view, the data suggests performance with a physical controller is better than that achievable with a virtual hand controller. Which result, if any, are we to deem as valid?

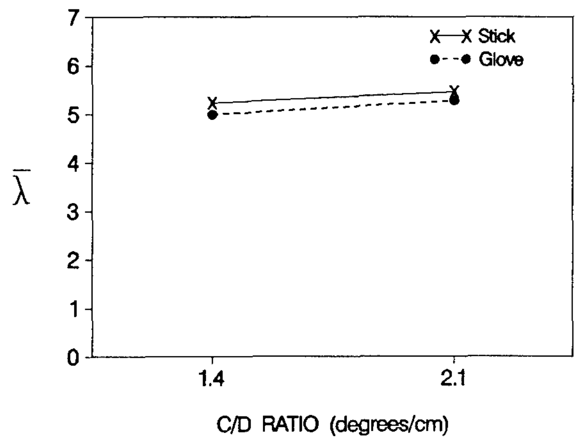


Figure 5. Comparison of Control Devices on the Basis of a Physical Identity Matching Scheme (See text for details.)

We are inclined to accept the results from the Performance Based matching method as most reflective of comparative user performance between the virtual GLOVE and physical STICK. Our argument is based on the following reasoning. First, the differences in kinematic and force demands of each control device are likely to interact with C/D ratio in determining user performance. Differences in user performance have been found with return-to-center and non-return displacement sticks. While the resistive forces for the GLOVE probably fall in between these two types of physical devices, it is still likely on this basis alone that optimal performance will not occur at the same C/D ratio as with the STICK. Second, the TLX workload data indicated that there was no perceived difference in effort required to perform with either hand controller at any of the C/D ratios used in the study. This suggests the subjects did not compensate for a harder task by working harder. If such a compensation

occurred, then matching devices on the basis of performance alone would not be adequate. The fact that the subjects apparently did not compensate in this manner means that in this instance the Performance Based matching method is valid without added correction for workload differences. Together these factors argue for using the results contingent on the Performance Based matching method. As noted earlier, these results provide convincing evidence in support of the claim that comparable control performance can be achieved with a virtual hand controller.

This study was limited to providing baseline information on user performance with a virtual control device. It did not investigate possible advantages of a virtual controller. Some advantages readily come to mind. A physically mounted controller is affixed at a given metric distance away from all users. A virtual controller may be placed where it is comfortable for the user. In other words, controller location is determined by the user in appropriate body coordinate space. This may be an important feature for future controllers that have to accommodate 5th percentile females through 95th percentile male pilots. One important question is whether or not a performance gain results from this scheme. Other possible advantages include: consistent/direct rotation and translation mapping to the desired output action, use as a back up system (it is always available), and in situ C/D rescaling to accommodate arm/hand injury. Future experiments are planned to investigate some of these ideas.

Emerging technology supports the development of virtual interfaces which may be approached from two different design perspectives: interfaces as virtual environments where the user is an occupant interacting directly with virtual tools, instruments, machines and vehicles; and interfaces as unique input/output control and display devices. Important questions of user performance are raised from both perspectives. It is imperative that user performance studies be performed now so that we will have quantitative human factors data needed to both guide and evaluate virtual interface concepts as they emerge with technological advances.

This experiment takes a tiny step toward the goal of producing a useful performance data base for virtual environment interfaces. It has demonstrated that considerable care is required in the analysis process to ensure unbiased comparisons are made between virtual and non-virtual alternatives. Finally, it provides solid evidence that comparable manual control performance can be achieved with a virtual hand controller in a single axis, compensatory tracking task framework.

REFERENCES

1. Boff, K.R. & Lincoln, J.(Eds.)(1988) *Engineering data compendium: Human perception and performance.* Vol. III, Section 12: control interfaces. Wright-Patterson AFB, Ohio: Armstrong Aerospace Medical Research Laboratory.

2. Chapanis, A. (1972) Design of controls. In H. van Cott & R. Kinkade (Eds.) Human engineering guide to equipment design. Washington, DC: US Government Printing Office.

3. McDonnell, J.D. & Jex, H.R. (1967) *A "critical" tracking task for man-machine research related to operator's effective delay time. Part II: Experimental effects of system input spectra, control stick stiffness, and control element order.* Systems Technology, Inc. Report No. NASA-CR-674.

4. Jex, H.R. (1989) *The critical-instability tracking task: Its background, development, and application.* Systems Technology, Inc. Paper No. 334.

5. Jex, H.R., McDonnell, J.D. & Phatak, A.V. (1966) *A "critical" tracking task for man-machine research related to operator's effective delay time. Part I: Theory and experiments with a first-order divergent controlled element.* Systems Technology, Inc. Report No. NASA-CR-616.

6. Jex, H.R. & Allen, R.W. (1970) Research on a new human dynamic response test battery. *Proceedings of the 6th Annual Conference on Manual Control,* April 1970, Wright-Patterson AFB, Ohio.

7. Wickens, C.D. (1984) *Engineering psychology and human performance.* Columbus, Ohio: Charles Merrill.

8. Wickens, C.D. (1986) Effects of control dynamics on perception. In K.R. Boff , L. Kaufman & J.P. Thomas (Eds.) *Handbook of perception and human performance, Vol. II: Cognitive processes and performance.* New York: Wiley.

9. Hart, S.G. & Staveland, L.E. (1990) Development of a multi-dimensional workload rating scale: Results of empirical and theoretical research. In P.A. Hancock & Meshkati (Eds.) *Human mental workload.* Amsterdam: Elsevier.

A NON-INTRUSIVE WAY TO MEASURE POINT OF GAZE

by

G.D.R. Zon and H.A. Mooij
Mooij & Associates
9 Leidsestraatweg
2341 GR Oegstgeest, NL

J. Bouwens,
Solutions And Projects
198 Turfschip
1186 XS Amstelveen, NL

SUMMARY

OBSERVER is an instrument for obtaining data about where a subject is looking on fixed user specified surfaces. Since the processing of data takes place in real time, this instrument can be used to indicate areas of interest just by looking at them.
In this paper, after an introduction on the application of point-of-gaze (POG) data, the OBSERVER system is described. Attention is given to subsystems as well as to calibration.
As the first application of OBSERVER, that of a measuring instrument, an "eye-witness quality experiment" is discussed.

1 INTRODUCTION

As control of systems is shifting more and more from direct manual control into the direction of supervisory control, the importance of a complete understanding of the cognitive factors involved in the human-computer dialogue becomes essential.

Studies of these cognitive factors involved in new Man-Machine Interface (MMI) subsystems are required in order to provide for safe, comfortable and effective operation.

Availability of information on visual fixation data (or "stationary point of gaze") could play an important role in the above studies. Only a limited amount of literature is available on the way visual fixation data should be used in support of such studies.

At the request of the European Space Agency (ESA), Mooij & Associates started developing the Point-Of-Gaze Measuring & Designation System mid-1991. The primary purpose was to produce a tool to measure and analyze visual fixation data of a crew member's point of gaze on any part of an MMI. In addition, the application of point-of-gaze data as computer input channel should be possible.

In Chapter 3, human body- and eye-movement sensors which form important components of a real-time point-of-gaze measuring system called OBSERVER will be described. Before that, however, the use of point-of-gaze data will be discussed in Chapter 2.

In preparation of the first application of the system in its originally intended role, a research project was carried out in mid-1993, in which OBSERVER was used to measure/record point-of-gaze information during an "eye-witness quality" experiment. This is the subject of Chapter 4 of this paper.

OBSERVER will first be used in a human-computer dialogue evaluation forming part of an MMI design study at the European Space and Technology Centre (ESTEC) during the fourth quarter of 1993.

2 THE USE OF POG INFORMATION

2.1 General

In literature on visual perception, temporal-spatial patterning and the duration of fixations are regarded as a reflection of the perceptual strategy used by an observer to extract meaningful information from a display.

The duration of a fixation period implies the relative importance of the display area to the observer, and is commonly interpreted by researchers as a measure of covert cognitive processing.
Research and development evaluations have been reported in which eye trackers have been used in the analysis of perceptual motor tasks, such as driving a

Presented at an AGARD Meeting on 'Virtual Interfaces: Research and Applications', October 1993.

car, flying an aircraft and certain sports activities. The method of data gathering in most cases was by recording the images of a "scene camera" with an overlay of a cursor indicating the measured eye line of gaze. The large disadvantage of this technique is that frame-by-frame digitizing after the tests is required in order to enter point-of-gaze information into a computer for analysis of the data. OBSERVER eliminates this drawback, since it enables automated point-of-gaze determination.
The following section discusses two areas of application of a measuring tool for a crew member's point of gaze on any part of an MMI. For practical purposes, (e.g. the mass of data) only systems with digital output can be used with success in these cases.

2.2 Point-of-Gaze Data for Evaluations

Mission-Planning System for Tactical Aircraft

The development of computer-processing power in recent years has led to an emphasis on factors such as speed of task execution and speed of data access in systems like those used in mission planning for tactical aircraft. The importance of rapid completion of the entire planning task, or part of it, is obvious.

The development of these systems has not primarily been driven by crew requirements, but rather by the developing technology. A consequence of this design approach is that, although technology has improved some aspects of the mission planning process, full advantage has not been taken of the technological advances available.
A better approach to the design of mission-planning systems would be to have the design of the interface (MMI) become a driving factor, since this area is becoming the main constraint with respect to improving the effectiveness of mission-planning, Reference 1. The design should be based on the crew-task relationship rather than on the task alone. A major implication of this approach is that the choice of technology may be derived from the definition of the man-machine interface. Point-of-gaze data measured and recorded during the evaluation by trained pilots of various options for MMI of mission planning systems could be beneficial. Statistical analysis of point-of-gaze fixation data would be an important post-processing activity.

Control of Systems and Experiments in Spacecraft

As more computer support and control are introduced, the operation of systems by crew on board spacecraft is changing drastically. There are very few flight opportunities, which means there is also little opportunity for evolutionary development of new systems or building up confidence through regular use.

In the future space station, the crew has a variety of tasks in both spacecraft system control and on-board experiment control. It is known that on-board crew time as well as ground-based crew time (for training) will be limited, resulting in conflicting requirements: the crew being involved in ever more activities while having ever less time available for training on each particular function, Reference 2.

ESA has contracted Mooij & Associates to employ OBSERVER in the evaluation of competing designs for user interfaces for controlling systems and experiments on board spacecraft.

2.3 Point-of-Gaze Data for System Control

In searching for new and better interfaces between systems and their users, it can be very useful to exploit an additional mode of communication between the two parties. Typical human-computer dialogues are rather biased in the direction of communication **from** the computer **to** the user. Animated graphical displays, for example, can rapidly communicate large quantities of data, but the inverse communication channel has a very low bandwidth. The availability of an additional, rapid information channel **from** the user **to** the computer would be helpful, particularly if it requires little effort on the part of the user.
Since OBSERVER measures a user's point of gaze - his focal point on a given surface - and reports it in real time, it is possible to use point of gaze as an input medium in user-computer interaction, and thus for system control.

Following are some arguments in favour of using point of gaze, amongst other, as a computer input channel:

- Point of gaze has a high bandwidth due to the fact that eye muscles, being extremely fast, are able to respond more quickly than most other muscles.

- Point of gaze, based primarily on eye motion, can be beneficial under high-g loading (eye motion under high-g loading is perfectly feasible).

- Shifting point of gaze comes naturally and requires no conscious effort.

The above arguments demonstrate that point of gaze is a potentially useful additional user-computer input channel, especially in situations where the user is already heavily burdened.

People do not normally move their eyes - and thus their point of gaze - in the same slow and deliberate way as when they manually operate computer input devices. The eyes continually dart from point to point in rapid and sudden saccades. Unfiltered point of gaze, therefore, cannot simply be used to replace computer input devices such as the mouse. This is why point-of-gaze fixations should be used, rather than unfiltered point-of-gaze data. The success of system control using point-of-gaze fixation data from OBSERVER has already been demonstrated.

3 DESCRIPTION OF OBSERVER

3.1 System

OBSERVER is a system with which a crew member's point of gaze on any part of an MMI can be determined and recorded. The recorded data can be used for statistical analysis. The system consists of an eye-tracking subsystem, a motion-tracking subsystem, a calibration and preprocessing subsystem and a work-station subsystem. Figure 1 depicts OBSERVER in the form of a block diagram. The system is capable of providing point-of-gaze data in real time. This characteristic makes it possible to use OBSERVER as a high-bandwidth designation tool (information **from** the user **to** the computer.

[OBSERVER incorporates simultaneous position-tracking of both hands. Being identical to head tracking, hand tracking will not be discussed further in this paper.]

During the design of the system, a lot of attention has been devoted to user-friendly calibration features. A description of the system is given below.

3.2 Subsystems

There are three essential subsystems: the Eye-tracking subsystem, the Motion-tracking subsystem and the Calibration and Preprocessing subsystem. In addition, the optional Work-station subsystem may be used.

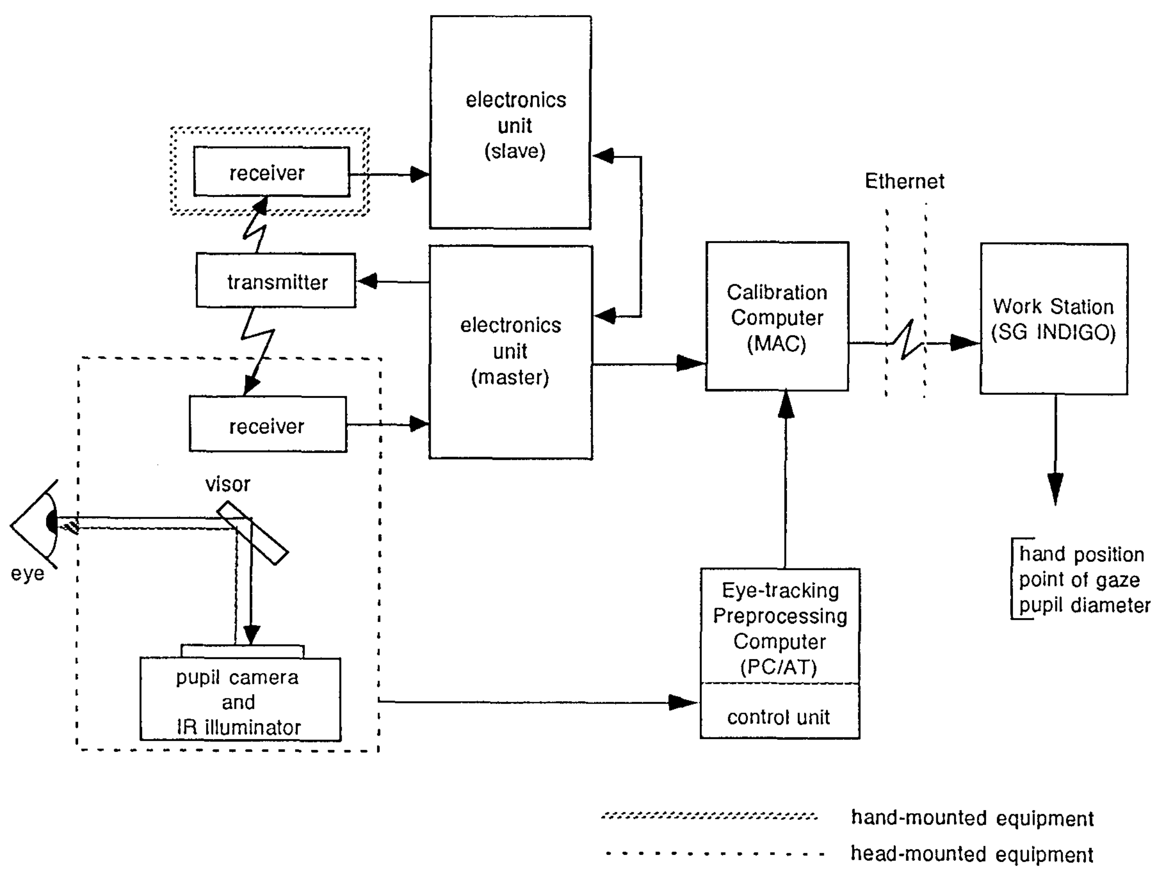


Fig. 1 OBSERVER

Eye-tracking subsystem

From the very limited number of systems suitable for incorporation into a "real-time" system, without head-motion restrictions, the Series 4000 Eye Tracker of Applied Science Laboratories (ASL) was selected. The mass and inertia of the head unit featuring no peripheral vision restrictions is of a level that prolonged wear is possible, the tracking range of eye line-of-gaze is an acceptable 50(H)x40(V) degrees, while the update rate is 50 samples/second. Eye-calibration time is short while the accuracy of 1 deg (RMS) is adequate for the application at hand.

The technique used is the pupil-to-corneal reflex vector method. The "bright pupil" version of the tracker was selected..

The system is controlled through a Subsystem Control Unit and a dedicated 486 PC/Monitor combination. Figure 2 presents a photograph of the optics module/visor combination.

During calibration of the eye, an "eye monitor" is used as well as a head-mounted calibration card.

Fig. 2 OPTICS MODULE/VISOR COMBINATION

Motion-tracking subsystem

The "magnetic type" position and orientation measuring system (indicated here as motion-tracking subsystem) of Ascension Technology Corporation (The Flock Of Birds) consists of a transmitter and a receiver both attached through cables to an electronics unit. The transmitter is the fixed reference against which the receiver measurements are made, while in this particular application the receiver is attached to a light-weight headband. The system works on the basis of a pulsed DC magnetic field. "Mapping" of the environment is not required. The position and orientation of the receiver anywhere within a sphere of 0.9 m radius is measured with an accuracy of 0.3 cm RMS for the position and 0.5 deg RMS for the orientation. The system has a maximum update rate of 100 samples/second.

Calibration and preprocessing subsystem

The calibration subsystem consists of an Apple Macintosh computer/monitor and the EPOG programme.
The function of the calibration subsystem is fourfold:

- It features three driver modules for communication with the eye-tracking subsystem, the motion tracking subsystem and the network driver (Ethernet).
- It provides a means to enter data during calibration. In this phase, the programme determines the position of the subject's eye in the system of coordinates of the (head-mounted) magnetic receiver. At the same time, it determines the scaling of the eye data to be used in the calculation of the point of gaze.
- It performs all required actions related to calibration of the system and measuring/-recording point-of-gaze fixations. Alternatively, measuring/-recording of point-of-gaze fixations may be remote-controlled from the Work station (see below).
- It facilitates the selection of certain parameters in the software, e.g. temporal and angular thresholds in the calculation of point-of-gaze fixations.

All four functions mentioned above are selected and controlled by means of a Dynamical Graphical User Interface (DGUI). All commands, selections etc. are given through a mouse (or tracker ball). Only the names of files and system settings are entered through the keyboard.

Work-station subsystem

The work-station subsystem is optional and not required when using OBSERVER for point-of-gaze system control.

The function of the work station is to store a high volume of measured data and provide a powerful platform for data analysis and visualization at a remote location. The work-station subsystem consists of a computer/monitor and the EPOGClient software programme. This programme facilitates recording of point-of-gaze fixation data in the work-station subsystem. The present target work-station computer is a Silicon Graphics Indigo, although in principle any system, computer and suchlike may be coupled to OBSERVER via Ethernet.

3.3 Calibration

To calibrate the eye tracker for use with a particular person, a short routine is performed during which data are loaded while this person alternately looks at nine points on a head-mounted calibration card.
To be able to determine the vector describing the receiver-to-eye separation, a short calibration routine is executed by the person under guidance of a test director. In this routine, the position of surfaces with respect to a room-fixed reference system is established as well.

Several tools have been developed for use in the calibration routine (a **stylus** temporarily to be attached to the magnetic receiver and an **eye-line bracket** attached to the "room" incorporating a "sight", a "receiver docking device" and a "receiver dummy block").

3.4 System Output

The pre-processed data related to point-of-gaze fixations are:

- Starting time of fixation
- Duration of fixation
- Surface identification
- X and Y of fixation
- Pupil diameter
- Distance eye to surface

The update rate of real-time point-of-gaze fixations is 20 ms.

4 EYE-WITNESS QUALITY EXPERIMENT

4.1 General

Following the above description of the system, a psychological experiment will now be described which was conducted with OBSERVER.

The problem discussed in this document concerns the fact that research into what subjects tend to remember about certain situations usually consists of asking them questions pertaining to the various stimuli to which they have been exposed during the test. As a rule, this is a satisfactory way of obtaining an impression of what people have remembered. A relevant example of this is an experiment taken by Wagenaar and Boer, Reference 3, the procedure of which was as follows:

Test subjects were shown a number of slides of aquarelles (phase 1). The slides told the story of a man leaving his home by car and a woman coming out of a shop. Their paths crossed literally on a pedestrian crossing which resulted in an accident.

One of the slides showed the man wanting to turn right at a junction while the traffic light was red (turning right on a red light is not allowed in the Netherlands).

Having carried out a diverting task, one half of the subjects was asked whether there had been a pedestrian crossing the street when the man wanted to turn right at the traffic lights. The other half of them was asked the same question, with the exception that the term 'traffic lights' was replaced by 'stop sign'. This is misleading information after exposure to the slide (phase 2). When subsequently shown two slides, i.e. the original one and another in which the traffic lights had been replaced by a stop sign, and asked which of the two they had seen, a number of the subjects selected the second slide (with the stop sign) (phase 3). The experiment entailed more, but this will suffice for the purpose of introducing the following.

After reading the Wagenaar and Boer's (Ref. 3) article on this experiment and having been made aware of the possibilities offered by the OBSERVER, it seems reasonable to assume that there must be a way of determining what test subjects actually see while looking at the slides. Information like that may well be a useful supplement to the above experiment, since it will provide not only answers to the questions posed, but also tell us at what the subjects were looking or, at any rate, what details they were looking at for any length of time.

In the experiment described here, the same slides were used as the ones used by Wagenaar and Boer (Ref. 3). Questions were made to go with each one of the slides, while for the specific slide with the traffic light, the questions described above were used. Subsequently, black and white photocopies were made of all slides. For each photocopy made, another was made in which one detail had been altered, for example, traffic lights altered into a stop sign. The slides were presented to the subjects as before, only this time they were monitored by the OBSERVER. For two slides it was recorded at what details the subjects had been looking.

Ultimately, it was possible to determine the relation between the amount of time they had spent on a certain detail and their selection of the photocopy of the original slide or the modified one. This way it became possible to establish an indirect measure to compare the amount of consideration (i.e. time) subjects had given to a certain detail, with their ability to correctly or incorrectly interpret the drawings in phase 3. In this article, time is considered to be an indirect way of measuring the attention a subject has for a certain detail. Time, therefore, is the unit in which the amount of attention is expressed.

The hypothesis was that subjects who spent little time examining a certain detail (e.g. traffic lights) would be more likely to be influenced by the misleading question than would the subjects who examined this detail at length. The underlying thought for this hypothesis being that the first group of subjects was only adding information to their memory, while the second group had to alter (already stored) information. This, it was felt, was less likely to happen. For a comprehensive description of the investigation performed, please turn to Reference 4.

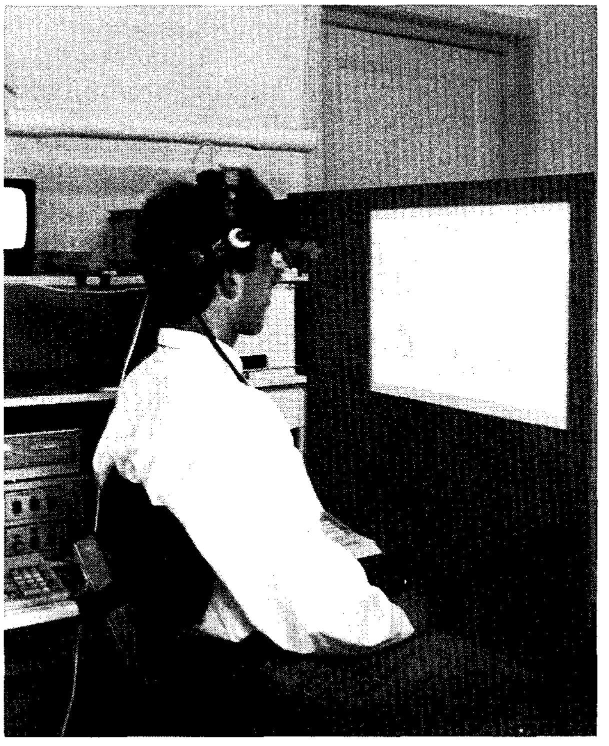


Fig. 3 SUBJECT OBSERVING SCREEN

4.2 Setup of the Experiment

Test subjects

The test subjects, a total of 35, were all students at the State University of Leiden, in the Netherlands.

Nine of the 35 subjects tested could not be used for analysis, since there was reason to believe that registration of their eye data was not up to standard. This type of problem sometimes occurs with persons wearing certain types of contact lenses which may distort the image recorded by the camera and used by the computer to determine the position of the pupil in the eye.

Experiment

First calibration was performed, as described in paragraph 3.3. Once this was completed, the actual experiment could start. The test subject was placed right in front of the slide screen as depicted in Figure 3. Twenty-one slides were used which together formed a story. Two of the slides were used to record what the subject had been looking at.

After having studied the slides, the subjects were given a separate task in order to divert their minds. This task took about eight minutes and had no bearing on the experiment at hand. The task consisted of having to look at colours on a computer screen.

Following the diverting task, they were given a list of questions regarding each of the slides with a choice of two answers per question (yes/no). Questions regarding each of two slides - with respect to which eye data had been recorded - came in two versions: a neutral one and a misleading one. The versions were distributed in a random fashion so that each subject had one neutral and one misleading question. Apart from that, the questionnaires were so composed that the number of questions to be answered with 'yes' equalled the number of those to be answered with 'no'. Questions regarding the other slides were used to mask the fact that some of the questions were deliberately misleading, so that subjects would not be aware that they were being manipulated.

Following the questionnaire they were shown photocopies referred to in paragraph 4.1. Of each pair the subject had to indicate which of the two versions he thought he had seen.

4.3 Results

Before analysis of the data could be started, there were some decisions to be made. In literature on this subject, there is no such thing as an exactly defined central field of vision. Since the borderline between periphery and centre is vague, the values are a matter of interpretation. For analysis, however, it is necessary to establish what a subject has actually

seen. One thing is sure, what is seen by the subject is more than just the point of gaze as calculated by the computer. The fovea (the area of the retina which has the sharpest vision) is considered by some authors as the central point of vision. Nonetheless, experiments show that there is a decline, however gradual, in the amount of detail seen from the centre of the fovea to the periphery and stored in memory, while lots of details away from the centre of vision are noticed by the subjects as well. In the end, it was estimated that the central area of vision of our subjects to be an approximate circle with a 3 cm radius on the screen.

The deviation from the point at which the subject was actually looking was calculated in relation to the point where the computer calculated the subject was looking. This was possible, due to the fact that subjects were asked after calibration to fixate on a cross in the centre of the slide screen. Since the exact position of this cross of course was known, it was possible to compare this with the point of gaze calculated by the computer. The circle described above was used to determine whether the subject had seen a certain detail in his/her central area of vision or not.

Leaving out subjects with doubtful eye data, there remained 26 subjects who were asked 2 questions each. In the case of the first slide, 11 of the subjects given neutral questions chose the correct picture and only two were inaccurate. Of the subjects given misleading questions, only 8 were able to make the right decision while 5 were mistaken. When using a Fisher exact test, this difference is not significant.

In the case of the second slide, however, 11 subjects given neutral questions made the right decision and zero were mistaken. Of the subjects given misleading questions, 9 were able to make the right decision and 6 were mistaken. When using the Fisher exact test again, there is a significant difference between the two groups at a level of 5%.

When considering the two slides per subject as 52 (26+26) reactions, the result is the following. Of the subjects with neutral questions, 22 chose the correct picture and only two were inaccurate. Of the subjects given misleading questions, only 17 were able to make the right decision, while 11 were mistaken. This means that there is a significant difference between the two groups. It may be concluded, therefore, that misleading questions indeed have an effect on the subjects.

This having been established, it is possible to analyze the OBSERVER data regarding the subjects who had to deal with a misleading question. The important contribution of OBSERVER is the way in which it is possible to look at the difference between subjects who made the wrong decisions, and those who made the right decisions. It would appear that subjects who had given little attention to the relevant detail on the slide and subsequently had to deal with misleading questions, were more likely to be "fooled" by the misleading questions than those who had spent some time examining the detail, e.g. the traffic lights.

In the case of the first slide, the average time spent on looking at the relevant detail by subjects who were given a misleading question and who were inaccurate, was 210 ms, while subjects who were correct took 730 ms. Using an analysis of variance, this is a significant difference ($F(1,11) = 5.8$ with $p < 0.05$).

In the case of the second slide, these figures were 353 ms and 492 ms, respectively, which is no significant difference. Combining all slides to 52 responses, as was done above, the result is as follows: the average time spent on looking at the detail by the subjects who were inaccurate was 287 ms, while the subjects who were correct looked at the details for about 604 ms, see Figure 4. Using an analysis of variance, this difference is significant ($F(1,26) = 5.4$ with $p < 0.05$). So the time a subject spent on a detail about which misleading questions were asked, tells us something about his knowledge of this detail. This means that the misleading question did not actually change information already stored in the subject's memory, but merely added information about something of which the subject had no knowledge at all. So, from a psychological point of view it is interesting to notice that OBSERVER makes it possible to gain a better insight into a subject's "viewing behaviour" than was possible before.

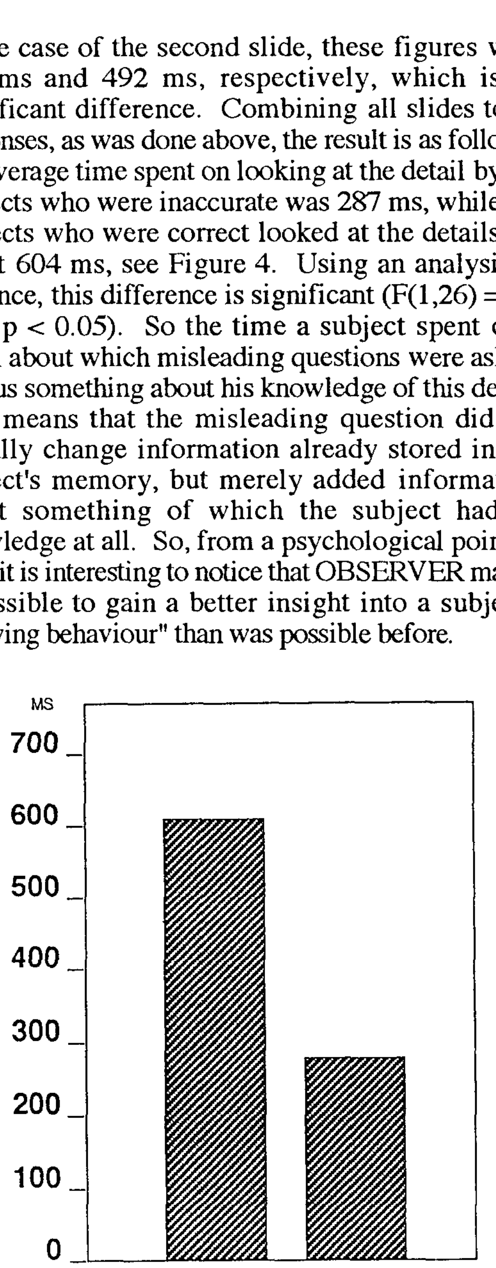


Fig. 4 AVERAGE LOOKING TIME

In conclusion, it could be said that the point of gaze data generated by OBSERVER may be used as a way to indirectly measure a subject's attention. The ordinary way of measuring that which is stored in memory is by asking the subject questions about a certain event or situation with which he is familiar. And, this information may be reliable if subjects are not influenced after - or even before - exposure to that event or situation. When subjects have not given any attention to the details about which they are questioned, they tend to be easily influenced by a misleading question. OBSERVER represents a less influenceable and more accurate way of measuring attention in (experimental) situations.

5 CONCLUSION

OBSERVER is a user-friendly measuring device, delivering point-of-gaze fixation data without interfering with a person wearing the light-weight headband holding optics, visor and a miniature magnetic receiver.

It has been shown that the Series 4000 Eye Tracker of ASL and The Flock Of Birds position and orientation measuring system of Ascension are robust sensors which performed well in the very first experiment with the system ("eye-witness quality experiment").

On the basis of the experience gained with OBSERVER, it is expected that point-of-gaze fixation data can be very helpful in MMI development.

6 REFERENCES

1. "Mission Planning Systems for Tactical Aircraft (Pre-Flight and In-Flight)", AGARD Advisory Report 313, 1992.

2. Ockels, W.J., "A New Approach to Spacecraft Crew System Operations", ESA Journal 1992, Vol. 16, pp. 233-236.

3. Wagenaar, W.A. and Boer, J.P.A. "Misleading Post-Event Information: Testing Parametrized Models of Integration in Memory", Acta Psychologica, 66, pp. 291-306.

4. Zon, G.D.R. "The Use of Point-Of-Gaze Measurement in Memory Research", Thesis October, 1993.

Operator Gaze Position Control Interfaces: Investigation of Psychophysical and Operational Parameters

Dave M. Stampe
Eyal M. Reingold
SR Research
55 Oakmount Rd. Ste. 1108
Toronto, Canada M6P 2M5
dstampe@psych.toronto.edu
reingold@psych.toronto.edu

Julius J. Grodski
DCIEM, Toronto, Canada

SUMMARY

Real-time monitoring of an operator's gaze position on a computer display of response options may form an important element of future computer interfaces and teleoperation control systems. In one implementation, the gaze position can serve as a pointer, and a critical length of gaze serves as selection. leaving the operator's hands free for other tasks. Control tasks such as multiple option selection, or looking for targets embedded within a picture are especially suited to selection by gaze position monitoring, since the search usually terminates on the object to be selected. More complex control functions can be implemented through multilevel "menus" of choices.

In the past, gaze monitoring systems restricted operator movement or required head restraints. The newest generation of gaze tracking systems allow free head movement and accurate gaze position monitoring over extended periods and are highly suited for control applications. Although gaze position control systems have been tried with moderate success in the past, little systematic investigation of the human parameters of gaze position control has been carried out. In the present research program, important parameters of gaze selection such as fixation position accuracy, selection error rates, and the effects of real-time gaze position feedback were investigated. Experimental results will be used to suggest guidelines for creation and use of gaze position response in control interfaces.

1. Introduction

The control of computers and other devices by monitoring an operator's gaze--"control by looking"--is becoming increasingly feasible for wide use as an interface method. Eye tracking equipment has become less expensive and intrusive, and computers and software needed to process eye position for control are more affordable. But the most important component in the control process--the user--remains little understood by most developers of gaze control systems. In this paper, we will develop some guidelines and methods for gaze control systems, based on our experience with the psychophysics of human eye movements and gaze control. We will also describe processing techniques used in a control system of our design, and some results of experimental tasks used to verify the system.

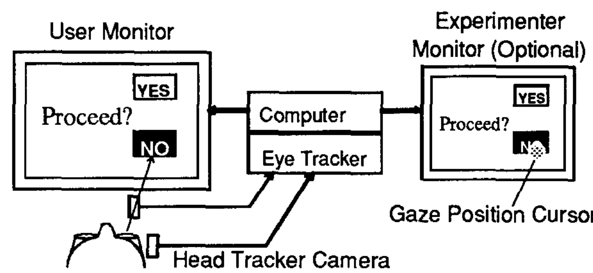


Figure 1. A typical gaze response system. The user's eye movements are monitored by an eye camera and tracking device, and a computer computes gaze position and duration. Sufficient gaze duration on the "YES" or "NO" response areas registers an appropriate response, and feedback is given to the user by highlighting the selected response. The second monitor is used in supervised tasks by the experimenter to monitor gaze position.

The use of control by gaze as an aid to handicapped persons is not new [1], but such systems have usually been developed with little or no research into psychophysical or cognitive factors. Gaze-controlled weapons aiming systems [2] have been investigated with moderate success, as has its use in telerobotics [3]. Recently, eye movements have been proposed as computer interface devices for normal users [4], but despite high expectations, the advantages and limitations of this interface modality have yet to be elucidated. There are many tasks in which gaze control would be advantageous, for example as an input device when hands are involved in other control tasks.

In a typical implementation, the user's gaze position is monitored by an eye-tracking device while viewing a task display presented on a computer monitor as in Figure 1. The user controls the system by directing his gaze to response areas or targets on the screen, and holding his gaze until the command is registered by the system. The screen may show computer-generated targets in addition to live video from a telerobot or computer graphics, and gaze position on any of these can be interpreted as a command. The gaze-response system processes the eye tracker output in real time to compute gaze position on the screen and to detect command events, then modifies the image displayed to the user in response to the gaze input.

Presented at an AGARD Meeting on 'Virtual Interfaces: Research and Applications', October 1993.

2. Characteristics of Human Eye Movements

Studies of the spatial and temporal characteristics of human eye movements provide important information for the design of gaze control systems. Eye movement may be divided into two phases: fixations and saccades. Fixations are periods when the eye remains stationary, gathering visual data, and are typically 150 to 500 msec in duration. In a pursuit eye motion, the eye moves smoothly to track a slowly moving object: this is similar in function to a fixation. Saccades are rapid (up to 800° per second) motions of the eye from one fixation position to another.

Research has shown that the position of gaze during a fixation indicates the locus of visual attention. Gaze tends to be captured by objects in an image, and it is difficult to fixate blank areas on a screen intentionally. Recordings of gaze position also show that gaze may be directed to the center of a cluster of objects, maximizing the information derived from each fixation. Gaze is also captured by moving objects in a scene, or by new objects appearing. This capture can be extremely rapid: express saccades may begin less than 100 msec after appearance of an object.

In normal vision gaze position shifts rapidly, an average of four times per second. The study of the duration and order of fixation of objects, words or other gaze targets is an important tool of psychological research into the mechanisms of reading, problem solving, and perception. The duration of a fixation on an object corresponds to the amount of cognitive processing it requires, and fixations longer than 500 msec may be seen during difficult tasks. These temporal parameters will be important in the design of gaze control interfaces as well.

3. Human Factors of Gaze Control

To the user, gaze control feels quite natural, due to the close link between attention and gaze position. Response by holding gaze on a target until a response is registered is especially natural, and is ideally suited to search within pictures or response arrays. Once the search or response target (e.g. an item in a menu) is found all the user need do is to continue fixating it until the response is registered. This is a highly intuitive response method, requiring no training and resulting in fast reaction times. Targets can be selected even when embedded within pictures or a dense arrays of other objects.

Gaze control acts as an extension of the user's ability to manipulate the world. Although objects in the world (except other humans) do not usually respond to our gaze, users immediately accept this method of interaction without need for training. To maintain the user's belief in the link between gaze and response, the gaze control system must be interactive: it must rapidly and predictably translate commands in the form of eye movements into system responses. These responses may take the form of presenting the next trial in an experiment, highlighting the selected response, or carrying out a command such as moving a piece on a displayed game board. The feedback serves to inform the user that the command has been registered, and must be visible while gaze remains on the response target.

Predictability is essential for gaze control systems. If the user looks at a response target, it is disconcerting if the target next to it is selected. This may occur because the eye tracking system has miscalculated the gaze position due to system noise or drift. Performance is also degraded if the gaze time required to select a response is unpredictable or if selection does not occur at all. Such errors will cause the user to distrust the system and impair performance substantially. Careful layout of response areas on the screen, use of high quality eye tracking systems, drift correction, and use of the reliable response detection methods discussed below will prevent these problems.

Achieving rapid responses to gaze input requires an integrated system capable of real-time processing of data from the eye tracker into gaze position. Other processing required includes detection of saccades, fixation and blinks, generation of gaze response events, and updating the display for user feedback and information presentation. Delays between any user response (via button press, voice key, or eye movement) and resulting changes of the display must be short and predictable: real-time processing systems are designed to fulfill this requirement.

4. Implementation Issues

To make gaze response practical, psychophysical and implementation issues relating to response error rates, response recognition, and user feedback must be addressed. The need for real-time interactive system response, the effects of eye tracking accuracy on performance, and some methods for correcting gaze position drift will be discussed. Techniques for analyzing the eye movement data for gaze response events will be introduced and evaluated.

4.1 Eye Tracking Devices

The basis of any gaze control interface is the eye tracking device. Many techniques have been developed for monitoring eye position, but only a few have the characteristics required for reliable gaze control systems. The environmental requirements for the operation of the eye tracking systems must be considered as well: systems that require head restraints or total darkness are unlikely to meet human factors requirements.

The eye tracking device function is utilized to determine the user's gaze position in the display of response targets and information. The tracker itself typically measures the eye position in a video image of the eye, by pupil position and/or by a reflection from the cornea of the eye. This position must then be converted to a gaze position by a mathematical transformation, determined by a system calibration. The calibration measures several (typically 5 or 9) point correspondences between gaze position and eye tracker output to compute the coefficient of this function [5]. Each point is collected by displaying a target on the display device and recording the eye tracker's data as the subject fixates the target. The accuracy with which the target is fixated and the position of the eye for each point is measured will determine the accuracy to which gaze position can be measured later.

Movements of the user's head in relation to the display device will cause the relationship between eye position and

gaze position to change, causing errors in computed versus true gaze position. It is essential that some method of compensation for head position is included in the gaze control system, since restraining the user's head is not acceptable in most applications. One method mounts the eye camera above or below the display monitor and uses the relationship between the position of the pupil and a reflection from the cornea of the eye as an invariant measure of eye rotation [6]. Other systems use a head-mounted camera to view the eye and track head position optically or magnetically, integrating these data by software to compute true gaze position on the display.

4.2 Resolution and Accuracy

It is important for the eye tracking system to have good resolution, accuracy, and stability. Resolution is the smallest change in gaze position that can be sensed by the eye tracking device itself, and is set by the noise level of the system or by the pixel resolution of the eye camera used. If resolution is too low, the system will not calibrate well and will not be able to distinguish between nearby locations on the presentation screen. This will severely limit the number of response options that may be presented together on the display. Resolutions of commercial eye-tracking systems range from 0.01 degree of visual angle to 1 degree or larger for some video-based tracking systems.

Accuracy measures how well the true gaze position of the user corresponds to that computed by the system. It is highly dependent on the quality of system calibration, which depends on how accurately the user fixated the calibration targets, and on the eye tracking system resolution. We find it useful to check accuracy immediately following calibration by displaying a set of targets for the user to fixate, and observing the computed gaze position as displayed by a cursor. If fixation errors on any target are too large, calibration may be repeated immediately.

Poor accuracy can result in gaze position errors of several degrees, causing the user's computed gaze to miss response targets or to select neighboring targets in error. Increasing the distance between targets reduces selection errors, but reduces the number of responses available to the user. Some systems developed for use by the handicapped were limited to nine response areas on the screen in a 3 by 3 grid to achieve acceptable selection error rates [7]. The small number of possible responses necessitated multiple screens of selection menus for all but the simplest commands, and made tasks such as typing text by eye slow.

4.3 Stability and Drift

A gaze tracking system with good stability will retain its accuracy for long periods of time after calibration. The most common form of instability is drifting of the computed gaze position away from the real gaze position, and is caused by eye tracking device factors such as head movements, shifting of the eye camera, or illumination changes. Even systems that compensate for head movements or use corneal reflection to reduce sensitivity to eye camera motion require periodic correction for drift.

A brute force method to compensate for drift is to simply recalibrate the system when required. A much faster technique is to measure the drift at a single point on the screen, then to apply a corrective offset to gaze positions on the entire screen. To compute the correction, a single target is displayed and fixated by the user, and the computed gaze position subtracted from the real target position. This process of *recentering* may be done on demand or scheduled between control tasks, and dramatically improves stability [5].

In our experiments a secondary monitor was used to display gaze position in real time, allowing the experimenter to judge when recentering was required. In a more typical gaze-control system, the user will be operating the system by himself, and must initiate recentering when drift becomes unacceptable, ideally before selection errors occur. Judging when recentering is required needs some form of gaze position error feedback, such as display of a gaze position cursor [4]. However, we have found such cursors, even when made transparent, to be distracting and to interfere with task performance. Under certain conditions, such as when the cursor had a small offset from the true gaze position, the user becomes trapped into uncontrollable pursuit of the cursor. It may be possible to place the display of the cursor under user control to prevent such problems.

4.4 Dynamic Recentering

A drift correction technique which is invisible to the user was developed for our gaze control system, which dynamically estimates and corrects drift during system operation. For small targets, it may be assumed that the average gaze position during a response is at its center. Therefore the average offset between the target and gaze position during responses will be the offset between true and measured gaze positions. This drift usually accumulates slowly, and the average error of several target fixations is sufficient to produce a running estimate of the drift for correction of gaze position. This incremental technique dynamically performs the recentering operation to correct system drift at each gaze response event, and can be combined with normal recentering to correct for catastrophic large drifts. Large changes in gaze offset are also gradually reduced over several selections.

The dynamic recentering algorithm is implemented as a simple low-pass filter which tracks the drift component of target fixation error, ignoring small random differences in target fixation. The C code for such a filter is given in Listing 1.

```
/* Implementation of dynamic recentering        */
/* DIVISOR sets the cutoff of the lowpass filter */
/* values of 3 to 6 work best                    */

static Xdrift = 0;
static Ydrift = 0;

Xcorr = Xgaze - Xdrift;
Ycorr = Ygaze - Ydrift;
Xdrift = Xdrift + (Xcorr - Xtarget)/DIVISOR:
Ydrift = Ydrift + (Ycorr - Ytarget)/DIVISOR:
```

Listing 1. Code for implementation of dynamic recentering algorithm.

5. Investigation of Accuracy

An experiment was performed (described in detail in [8]) to investigate how accurately users could fixate small targets on a computer monitor. The results were analyzed to determine the efficacy of dynamic recentering in correcting eye tracker drift, and the effect of distance between response targets in determining selection errors. The variation in fixation position of targets was also computed.

The experiment was performed using a prototype SR Research eye tracking system. This system uses a headband-mounted eye camera and scene camera which views LEDs on the display monitor to track and compensate for changes in the user's head position. It has a resolution of better than 0.005° (15 seconds of arc) for eye tracking, and compensates for head motions to better than 1 degree of visual angle over ±30° of head motion. Eye position is sampled 60 times per second, and noise is extremely low.

Target displays were presented on a 21" Idek VGA monitor located 75 cm in front of the user. A second VGA monitor displayed the gaze position in real-time to the experimenter, and was used to perform calibrations and to verify accuracy. Gaze position accuracies of better than 0.5° on all parts of the screen were routinely obtained through this verification procedure. The software was run on a 486/33 IBM PC compatible computer, including the real-time gaze control system.

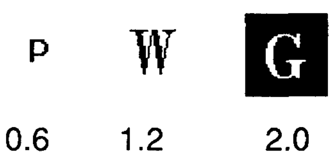


Figure 2. Targets used in the single-target accuracy tests. Size varied from 0.6° to 2.0° in visual angle.

5.1 Single Target Fixation

The first task was designed to investigate the effect of target size on accuracy of fixation. Targets may not be fixated centrally because of global target characteristics such as size or shape [9], and gaze may be more widely distributed on larger targets. It is important that targets be fixated centrally if dynamic recentering is to be used. In the task single targets subtending visual angles of 0.6°, 1.2°, and 2° as in Figure 2 were displayed sequentially in different positions on the monitor. The targets were to be fixated for 850 msec to register a selection, after which the next target was displayed.

The magnitude of the fixation error was computed as the Euclidean distance from true target position to the response gaze position. During analysis, dynamic recentering was simulated to correct gaze position for drift, and proved effective: mean fixation error was reduced from 0.73° before correction to 0.5° after correction. The small fixation errors were likely the result of psychophysical inaccuracies in fixation rather than equipment noise or inaccuracies, as system accuracy is known to be better then this. Target size did not influence fixation accuracy significantly, implying that single symmetrical targets subtending as much as 2° were fixated centrally in this task.

O	O	O	O	O	O	O	O
O	O	O	O	O	O	O	O
O	O	O	O	O	O	O	O
O	O	O	O	O	T	O	O
O	O	O	O	O	O	O	O
O	O	O	O	O	O	O	O
O	O	O	O	O	O	O	O
O	O	O	O	O	O	O	O

Figure 3. A sample target array used in the fixation accuracy tests. The "T" is the search target to be indicated by gaze response. This is an 8 by 8 array of response targets spaced by 2°.

5.2 Accuracy in Arrays

The second task explored the effect of the layout of arrays of multiple response targets on fixation and selection errors. Increasing the distance between targets should decrease the likelihood that an eye tracking or fixation error will cause a target adjacent to the intended one to be selected by mistake. Target spacings of 2° and 3° were used, as 4° spacings were known from previous trials to make the likelihood of selection errors vanishingly small. One dimensional (line) and two dimensional (grid) arrays were also investigated.

The response targets in the array were small characters: one "T" character to be indicated by gaze response, embedded in "O" distractor targets. The search target was highly salient to keep search times short and minimize errors. A typical search array is shown in Figure 3. A gaze dwell time of 1000 msec was required to select the target. Selection of an "O" distractor character was counted as a selection error.

Users reported that the task was not difficult, and that the 1000 msec gaze dwell time seemed appropriate. Selection errors were somewhat higher for two-dimensional target arrays (3.3%) than for one-dimensional line arrays (1.3%), as targets have more neighbors in the two-dimensional case and thus a greater likelihood that a fixation error or drift will cause a neighboring target to be selected.

The effects of drift and dynamic recentering are best shown by a histogram of fixation error magnitudes in Figure 4, computed as the distance from the center of the "T" response target to the response gaze position. Most of the fixation errors are less than 0.6° in magnitude. This probably represents psychophysical variations in fixation position on the target, as the tracking system accuracy is known to be better. The long tail in the uncorrected fixation error distribution is due to larger errors caused by eye tracker drift. This tail is significantly reduced by the application of dynamic recentering, which reduces the probability of large errors by a factor of 3.

As in the first task, dynamic recentering reduced the average fixation error only slightly, from 0.51° before

correction to 0.38° after correction. It markedly reduced selection errors from 6.6% before correction to 2.4% after correction. This clearly indicates the effectiveness of dynamic recentering in preventing the buildup of large gaze position measurement errors due to tracker drift. The probability that a fixation error will cause a selection error mat be computed from the histogram of Figure 4, by computing the area under the distribution curve for fixation errors greater than half of the distance between response targets. For example, at 2° target spacing, fixation errors greater than 1° will result in selection errors, and occur in 9.5% of trials without dynamic recentering, which occurred in 2.8% of trials after correction. For 3° target spacings, selection errors were seen in 1.9% of trials with dynamic recentering applied.

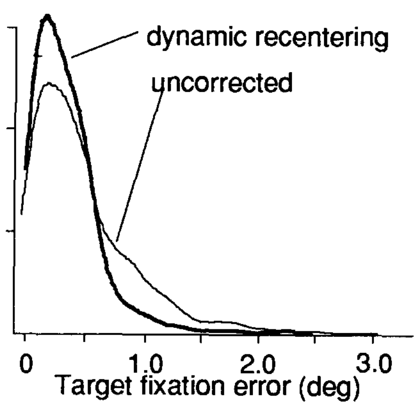


Figure 4. Distribution of fixation errors by magnitude before and after correction by dynamic recentering. Note the long error tail in the uncorrected error distribution caused by eye tracker drift. The incidence of fixation errors larger than 0.6° is greatly reduced by recentering, decreasing the incidence of selection errors.

6. Response Triggering

The basic mode of gaze control is similar to the operation of a computer mouse: gaze position is used to indicate a position or response option on a display, and some method analogous to a mouse button press is used to trigger the response. Holding gaze on the response target is the most intuitive method for selection, but relatively long (500 to 1000 msec) gaze dwell thresholds may be required to prevent unwanted responses from being registered. Other methods of registering responses include buttons, voice key, confirmation targets, or blinks.

Blinks can be detected by most eye tracking devices, and have been used in aids for the handicapped, often as a yes/no selection device. However, blinks disrupt gaze position data, and the long blinks required for control (as long as 500 msec) are tiring and can cause irritation to contact lens wearers. Blinks might be used by the user to request important events such as recentering or system calibrations since these do not require gaze position input.

A button press or triggering of a voice key could be used in combination with gaze position to select a response target. This response method can be faster than triggering by gaze time, especially if a long gaze dwell time is used. Unfortunately this method often results in selection errors, as the user often simply glances at the target while initiating the button press. Because eye movements are initiated more rapidly than the button press, the eye is often no longer on the target by the time the button press is completed. Training is required to reduce the occurrence of these anticipation errors. A long (500 msec or greater) fixation could be required in combination with the button press, but this does not have a great advantage over a long fixation alone. Button selection was tried in our initial studies of gaze control but found to be less successful than triggering by gaze duration, a result also found by Jacob [4] in his research on gaze control of computers.

In cases where a response has no irreversible effects or can be corrected immediately, short fixations of 300 msec or greater can be used to trigger control events. For example, Jacob [4] immediately responded to user's gaze on ship icons on a map by updating information in a side window containing information about that ship. When the window was looked at, it always contained information about the last ship fixated. This technique requires only short fixations on the targets, but is not practical when expensive or irreversible actions such as menu selection result from a control event. This method has been extended in a typing task [10] by requiring users to look at a confirmation target after fixating the desired character to be typed. Although this allowed short fixations to be used, the time required for the saccade and fixation on the second target made the speed increase marginal.

The most natural method of response is simply to hold gaze on the response area for a critical dwell time. Subjectively, the user simply concentrates on a target until the command is acknowledged. Most users can use this technique immediately, and need little or no practice to perform well. The duration of gaze needed to select a target must be short enough to be comfortable for the user, yet long enough to prevent unintentional triggering by long fixations. It is well known that the proportion of long fixations (greater than 500 msec) increases for difficult tasks. Longer dwell times may be needed for these tasks to reduce the probability of unintentional selections. Pilot studies indicated that a dwell time of 1000 msec makes false selections unlikely, and may be reduced to 700 msec or less for simple tasks.

6.1 Gaze Aggregation for Dwell Selection

Ideally, gaze on a response target would always be seen as a single long fixation. In our initial studies, gaze periods of 800 msec or more were often broken by blinks, corrective saccades and attentional lapses thus breaking up the gaze period into several fixations. If duration of single fixations is used to detect gaze time and trigger responses, system operation may be unreliable. Subjectively, the gaze time required to select the target becomes irregular due to broken fixations. If the total gaze time on the target exceeds 2 seconds, the user's gaze may shift involuntarily or the user may give up. Some users have reported severe eyestrain. This may account for some reports of problems with dwell times longer than 700 msec [4].

Detection of gaze by methods that allow aggregation across one or more fixations will prevent blinks or small shifts in gaze position from influencing the subjective response

delay of the gaze control system. Detecting fixations is the first step in the gaze aggregation, as blinks or tracker artifacts often appear as short or misplaced fixations. Fixations are defined as periods when eye position is stationary and are separated by saccades, which may be detected by rapidly changing eye position as described in [5]. Only data from fixations that exceed a time threshold of 80 msec (the shortest psychophysically valid fixation duration) are integrated into the gaze period.

After the fixation duration exceeds this validation threshold, its data must be integrated into the gaze estimate as each data sample arrives from the eye tracker. This allows the control response to be triggered at the instant that gaze time exceeds the dwell threshold, instead of at the end of the fixation. Position of gaze may be estimated by computing a running mean of the position data of all data points for all fixations in the gaze.

Gaze periods may be defined by detecting tightly clustered sequences of fixations. For arrays of response targets with well-defined positions, it is sufficient to define a region on the display for each target for gaze aggregation. Sequential fixations falling into a target's gaze region are aggregated to determine the gaze time. This simple method of aggregation may fail if the locus of gaze falls near the edge of a target's region, as some of the fixations in the gaze may fall outside the region and terminating the gaze period prematurely.

When targets are tightly packed or selection of targets is unstructured displays such as video images or pictures is required, a more robust method of gaze aggregation is needed. Cluster aggregation [11] integrates groups of fixations into a single gaze position, which can then be used to select a response target. The basic algorithm compares the position of each fixation to the center of gravity of a potential cluster, computed as the average of positions weighted by duration of each fixation. If the fixation is closer than a critical distance to the cluster, it is added to the cluster, otherwise the fixation begins a new gaze cluster.

For gaze control systems, a real-time variant of cluster aggregation is used. As each gaze position data point arrives from the eye tracker, it is integrated into the current fixation's average position, which is compared to the cluster's center of gravity. If the fixation is sufficiently close to the cluster, the position data is also integrated into the cluster's position. If the fixation's position deviates sufficiently from the cluster's position, the fixation's data is removed from the cluster and the fixation's data used to start a new cluster. When the number of samples in the gaze cluster exceeds the gaze response dwell time, the position of the cluster is used to determine what target or position on the display has been selected.

6.2 Integrated Gaze Control

To integrate gaze control with complex computer software, it must be encapsulated with a simple and standard communications method. The most usable paradigm is that used by mouse-driven GUI (graphical user interfaces) such as the operating system of the Macintosh or Windows. Here, the application software draws or creates interface elements such as menus or dialog boxes, and mouse clicks or other selection events are processed by the operating system and relayed back to the software as control events. It is clear that gaze control can emulate the mouse in these systems easily, although its feasibility to replace the mouse for applications such as drawing or text editing is not as clear.

In the implementation of such a system, processing of eye tracker data in gaze control systems would take place in real time, including calculation of gaze location on the screen and saccade and fixation detection and blink rejection. Gaze would be aggregated by cluster or region to produce current gaze position and duration, and responses would be triggered when gaze is in a valid target's response region and exceeds the dwell time threshold. Target gaze regions would be predefined by application or system software and used for fixation aggregation or to identify the selected target when cluster aggregation is used.

The response would be processed immediately to give user feedback such as proceeding to the next menu screen or highlighting of the selected response. If implemented, dynamic recentering would be performed automatically at the time of each selection. These operations will ensure stable and predictable responses. Control events would be passed back to application software for appropriate processing.

THE QUICK RED FOX JUMPED OVER THE LAZY BROWN DOG
THE_QUICK_RED_

A	E	I	O	U	<	□
B	C	D	F	G	H	J
K	L	M	N	P	Q	R
S	T	V	W	X	Y	Z

Figure 5. The screen layout of the typing task. All letter targets are separated by 4° to keep selection error rates low. The backspace and space targets are at the right end of the top row of targets. The top part of the screen contains the text to be typed as well as the typed output. Letters are highlighted for 300 msec as they are typed: the "F" is highlighted in this picture.

7. Sample Gaze-Control Tasks

Two gaze-control tasks are presented which investigate the feasibility of the interface method for practical uses. In the first task, typing by eye is implemented and compared to results by other researchers. In the second task, a board game requiring substantial mental effort is played to determine if gaze response and cognitively demanding activities are compatible.

7.1 Typing by Eye

Typing by eye is one of the most common tasks to be implemented in eye-movement control systems for the handicapped. It is unfortunate that little investigation has been done into its efficiency in the past, perhaps because of the focus on the implementation of such systems rather than on research. Typing requires at least the full alphabet as response alternatives, and demands high resolution and accuracy from gaze-control systems if all characters are to selectable at once. Where very few response alternatives per screen are available, multiple-level menu screens have been used [7], but resulted in relatively slow typing speeds.

The goal in this task was to evaluate the users' impressions and error rates during typing. The screen layout is shown in Figure 5, and used a 7 by 4 grid of 1.2° letters spaced by 4° horizontally and vertically. The top part of the screen contained the line of text to be typed and a space for display of the typed output. The users typed by fixating the desired letter for 750 msec, with gaze aggregated within a 4° square region centered on each target. Dynamic recentering was applied at each selection to correct for system drift. Selection feedback was given by placing a round highlight spot on the letter for 300 msec as in Figure 5. If the user continued to fixate the character, it was typed repeatedly.

Each user typed three text samples: the first was a practice trial where users typed their name or other random input. In the second trial, users typed the sentence "THE QUICK RED FOX JUMPED OVER THE LAZY BROWN DOG", a total of 48 characters. In the third trial, users typed " INTERACTING WITH COMPUTERS AND CONTROL BY EYE", 44 characters. Characters typed and use of the backspace were recorded along with fixation data.

Users found the typing method interesting, but slower than manual typing. The gaze dwell time of 750 msec subjectively seemed limiting especially during the last typing trial, but in reality typing speed did not increase during the experiment. The time spent in selection of each typed character was only 40% of the 1870 msec average time required to type each character. The remaining 1120 msec (an average of three fixations) was spent in searching for the character in the typing array. It is expected that search time will decrease with practice, permitting shorter dwell times to be used.

Errors were classified by counting backspaces and examining the typed output: transcription errors included missed characters or spelling mistakes (4 instances in 1400 characters typed). Selection errors were scored if a spelling mistake involved a letter adjacent to the correct letter on the selection grid, and occurred 5 times out of 1400 characters typed (0.36%). The low selection error rate for the 4° target spacing can be compared to the 1.9% error rate for 3° spacing and 3.4% for the 2° spacing measured in the search task discussed earlier. Error rates also compare well to the 1.3% reported for a 54-character typing screen [12], which can be expected to require much closer target spacing.

It is apparent that typing by eye is much slower than manual typing, with most time spent searching for the character to be typed. With much practice, search time may be minimized and the dwell time may be reduced further. Research has shown that typing by touch screen can be as fast as 500 msec per character (25 WPM), and by mouse at 700 msec (17 WPM) [13]. Typing by eye can probably be as fast once character positions in the typing array are memorized.

Pairs: 2 Flips: 17

1	2	hum	4	5
mom	con	8	9	10
11	hum	13	mom	15
16	17	18	19	20

Figure 6. The screen layout used in the "Concentration" game task. Squares are 4° in size and are numbered to encourage central fixation. Once fixated, squares "flip" to reveal a large letter or three-letter word. The number of flips and pairs found are displayed at the top left of the screen.

7.2 "Concentration" Game

It is important to investigate gaze control in conjunction with a more complex cognitive task, which increases the probability of very long fixations and therefore of unintentional responses. The game of "Concentration" was chosen, as it is easy to learn and to implement in software. In this game, players flip numbered squares to reveal letters or words. The objective is to reveal matching pairs of letters or words. If a pair of nonmatching letters or words are revealed, both are hidden again. This task requires a broad memory span and careful search strategy to minimize the number of flips required to reveal all pairs.

The game screen layout used is shown in Figure 6 and consisted of a 5 by 4 grid of 4° squares. As it is difficult to fixate the center of a large blank area, each square is numbered in its center to provide a target for gaze. Central fixation of the squares helps to reduce selection errors and improve dynamic recentering performance. The current score in total flips and pairs revealed was displayed at the upper left corner of the screen. A gaze duration of 1000 msec, aggregated within the region of the square, caused the small number to be replaced by a larger word or letter. This change was salient enough that no other selection feedback was required. If the flipped square was the second in a nonmatching pair, both tokens were hidden 600 msec later.

Users reported the method of play to be interesting and intuitive, although the game itself was not easy. An average of 43 flips and 100 seconds were required to complete each game. The dwell time of 1000 msec was judged to be correct for the task by all users. Even though many long fixations were seen due to the difficulty of the task were seen, none resulted in unwanted responses,

suggesting that a 1000 msec dwell time maybe sufficient for complex tasks.

Use of a shorter dwell time of 700 msec caused some subjects to report the subjective impression that the computer was anticipating their responses, flipping squares on the board before the user was aware of initiating the action. It is difficult to determine whether such anticipatory responses in fact match the responses that would have been made with to longer dwell times, or were simply unintended selections. Such anticipatory actions have been reported by other researchers when short dwell times are used [4].

8. Discussion

In considering the use of gaze control for any application, the human factors of the task must be considered carefully. Fine control of gaze position is not always possible: gaze tends to be attracted to visual objects, and is difficult to control on blank areas of the display. Adding extra detail to the display such as the numbers at the centers of squares in the board game task discussed earlier, or a grid of dots to help fixations be accurately placed on blank areas can be expected to improve task performance.

Eye movements have many automatic characteristics, and initial fixations on response targets may be influenced by global aspects of a display such as position and shape of targets [9]. Increasing gaze dwell times may also produce more accurate secondary fixations to improve fixation accuracy and decrease selection errors. Long dwell times may also result in more accurate task performance [8], as the dwell time allows the user to correct erroneous response choices. This increased accuracy must be considered against the increase in task performance time.

Some tasks may require careful design if they are to be used with gaze control. For example, if gaze control options are presented on an aircraft heads-up display, fixations on external objects may cause false selections. Gaze also may be stationary during driving, tending to rest on the center of visual flow. Long fixations to perform selections may impede operation of vehicles by preventing the operator from scanning the instruments or field of view. Using button selection of the fixated response option rather than gaze dwell time may be superior to gaze dwell time in such environments.

Accuracy of gaze on response targets was found to be very good, with mean fixation errors of 0.5° or better. Targets of up to 2° is size were fixated as accurately as smaller targets, but it is expected that very large targets (>4°) will not be fixated centrally, instead being fixated near an edge or corner. Adding central detail to such targets may be helpful. The greatest cause of selection errors was found to be drift, which dynamic recentering helped to reduce. Target spacings of 4° showed vanishingly small selection error rates, while target spacings of 2° or 3° showed low but significant error rates, largely due to drift. If an eye tracking system without drift was used, target spacing of 2° or less might be practical. However, it might be difficult to locate the desired response option in a large array of such tightly packed targets, degrading task performance.

Gaze control should be very useful in target designation tasks. In tasks where detailed image analysis is needed, gaze control may be used to concentrate image processing resources around the operator's point of gaze, substantially reducing processing requirements. Targeting of moving or stationary objects in a real scene by gaze may not be accurate enough for tracking or weapons control applications, due to psychophysical fixation inaccuracies or eye tracking resolution. It is possible to use gaze position to guide an image processing computer to search for exact target location in the image or to lock onto a moving target.

9. Conclusions

In general, the results from the experimental tasks suggest that gaze response is intuitive and reliable enough to be practical in many teleoperation and computer interface applications. All users performed a wide variety of control tasks without need for any training, and were enthusiastic about the natural quality of selection by looking. These positive subjective impressions were further supported by the speed and accuracy scores for the tasks.

Gaze control can be learned quickly, and feels natural to the user. However, gaze control is a simulated, nonphysical interaction method and depends on predictable and correct operation and prompt and visible response feedback to maintain the illusion of control. Users quickly become frustrated or hesitant if selection times become variable or response selections are incorrect. The real-time processing methods described in this paper help to improve system stability by making selection times independent of blinks or eye tracker artifacts, and reducing selection errors due to drift. These methods were verified in the experimental tasks described in this paper.

Important to the success of the paradigm was the ability to precisely place gaze on response targets and to hold the gaze for long enough to trigger the response. Although natural gaze is often broken by blinks or refixation, the aggregation of gaze by cluster or segment resulted in reliable selections and predictable gaze times. Users had no difficulty with dwell times requiring gaze periods as long as 1000 msec. This is in marked contrast to reported difficulty with dwell times over 700 msec by Jacob [4], who used only single fixation as a measure of gaze duration.

Psychophysical limits on accuracy of gaze placement were not large enough to be a problem in response selection. The main source of selection errors appeared to be the result of occasional drifts in the eye tracking system. Such drift could be corrected by the use of dynamic recentering, which reduced probability of selection errors by 66%.

If eye trackers with low resolution or with rapid drifting such as that caused by head movements are used, response targets must be widely separated, reducing the number of response options that may be placed on the display screen. Multiple-level menus of screens may be used to expand the number of options available, at the cost of increased selection time.

The typing and game-playing tasks were representative of gaze control computer interfaces. These tasks required reliable selection between many response targets, which mandates high-quality eye tracking devices with good accuracy and low drift. User comfort is important if gaze control is to be accepted by computer users, requiring

headband-mounted or desktop eye trackers that do not constrain head motion.

10. References

1. Laefsky, I.A. & Roemer, .R. A. (1978). A real-time control system for CAI and prosthesis. *Behavioral Research Methods & Instrumentation, 10(2)*, 182-185.
2. Abbey. (1968) Guide Weapons of Aircraft, *U.S. Patent 3,376,375*
3. Razden. R. and Kielsr, A. (1991). Eye Slaved Pointing Device for Teleoperator Control. Proceedings of the SPIE, v1388, pp361-371
4. Jacob, R.J.K. (1991). The use of eye movements in human-computer interaction techniques: What you look at is what you get. *ACM Transactions on Information Systems, 9(3)*, 152-169.
5. Stampe, D.M. (1993). Heuristic filtering and reliable calibration methods for video-based pupil-tracking systems. *Behavioral Research Methods, Instruments, & Computers, 25(2)*, 137-142.
6. Merchant, J. et al. (1974). Remote measurement of eye direction allowing subject motion over one cubic foot of space. *IEEE Transactions on Biomedical Engineering, 21*, 309-317
7. Frey, L.A, White, K.P. Jr., & Hutchinson, T.E. (1990). Eye-gaze word processing. *IEEE Transactions on Systems, Man, and Cybernetics. 20(4)*, 944-950.
8. Stampe, D.M. and Reingold, E.M. (1993). Eye Movements as a Response Modality in Psychological Reaserch. In J. Findlay (Ed.), *Eye Movement Research: Mechanisms, Processes, and Applications* (In Press)
9. Findlay, J.M. & Crawford, T.I. (1976). The visual control of saccadic eye movements: Evidence for limited plasticity. In R. Groner, C. Menz, D.F. Fisher, & R. Monty (Eds.), *Eye movements and psychological processes: International views* (pp. 115-127). Hillsdale, NJ: Lawrence Erlbaum Associates.
10. Charbonnier, C. and Masse, M. (1993) Typing Controlled by Ocular Command. Poster paper presented at the Seventh European Conference on Eye Movements, Durham, Sept. 1993
11. Kundel, H.L., Nodine, C.F., & Krupinski, E.A. (1989). Searching for lung nodules: Visual dwell indicates locations of false-positive and false-negative decisions. *Investigative Radiology, 24*, 472-478.
12. Spaepen, A.J. & Wouters, M. (1989). Using an eye-mark recorder for alternative communication. In A.M. Tjoa, H. Reiterer, and R. Wagner (Eds.), *Computers for Handicapped Persons*. (pp. 475-478). Vienna: R. Oldenbourg.
13. Andrew, S. (1991). Improving touchscreen keyboards: Design issues and a comparison with other devices. *Interacting with Computers, 3(3)*, 253-269.

Orientation du regard sous facteur de Charge
Aspects méthodologiques
Résultats préliminaires

Gaze Orientation under G_z-load
Methodological Aspects
Preliminary Results

Patrick B. SANDOR*
Isabelle HORTOLLAND*
Frédéric POUX*

*Centre d'Essais en Vol, LAboratoire de Médecine AéroSpatiale
Division Biodynamique, F-91228 Brétigny sur Orge, FRANCE

Alain LEGER **

**Sextant Avionique, Division Visualisations et Interfaces
Rue Toussaint Catros, F-33400 Le Haillan, FRANCE

Résumé :

La direction du regard d'un sujet soumis à une accélération G_z en centrifugeuse a été calculée en condition tête libre. Le mouvement de l'oeil était mesuré par un oculomètre du type pupille-reflet cornéen, le mouvement de la tête par un système électro-optique de détection de la position de casque. La tête du sujet était positionnée approximativement au centre d'un hémisphère de 1,80m de diamètre. La face interne de cet hémisphère constitue un écran sur lequel un spot laser est envoyé. La LDV du sujet est calculée à partir de la direction de l'oeil dans le repère mobile de la tête. Une procédure de correction d'erreur de parallaxe permet de calculer le point d'intersection de la LDV et de l'écran, déterminant ainsi les écarts cible-point de regard.

Après validation statique de la chaîne de mesure, deux expérimentations préliminaires sous facteur de charge ont été conduites. Les résultats obtenus démontrent la faisabilité de la méthode de désignation dans l'environnement expérimental. Les améliorations nécessaires à l'acquisition de données permettant une étude quantitative précise ont également été déterminées.

Summary :

Gaze in head-free condition was computed under G_z-load. Eye movements were measured with an oculometer using the pupil-to-corneal reflex method. Head movements were measured with an electro-optic system. The subject's head was at the centre of a hemisphere (diameter 1.80 m). The internal face of this hemisphere was forming a screen on which a laser spot was to be projected. The subject's line-of-sight (ligne-de-visée LDV) was computed, i.e. the direction of the eyeball in the head frame, which is mobile relative to the space. A procedure of correction of the parallax error allowed the determination of the Point-of-Gaze, which is the intersection point of the LDV with the screen.

After static validation, two pilot experiments were performed under low G_z-load. Results showed feasibility of the method in the experimental environment, an pursuit errors were quantified. Improvements are proposed.

Introduction :

Les environnements aéronautiques actuels des avions de chasse amènent progressivement une augmentation de la masse et souvent un déplacement du centrage des dispositifs portés sur la tête. Dans ce cadre, les systèmes de type viseur de casque pourraient bénéficier, en plus de la détection de la position de la tête, de la détection de la direction du regard (Ligne-De-Visée LDV) du pilote. Dans le futur, l'utilisation de la LDV permettrait l'amélioration de l'interface Homme-système, aussi bien pour les applications militaires que civiles. Dans le domaine militaire, les conditions de vol sous facteur de charge constituent un domaine d'emploi particulièrement intéressant.

L'intégration d'un système de mesure de mouvement de l'oeil et d'un système de mesure de mouvement de la tête a été entreprise au Laboratoire de Médecine Aérospatiale du Centre d'Essais en Vol depuis plusieurs années. Deux expérimentations successives menées sous accélération dans une nacelle de la centrifugeuse humaine du laboratoire. Elles étaient fondées sur l'erreur mesurée entre le point de regard, c'est-à-dire le point d'intersection de la LDV avec l'écran constituant le support de la cible, et la position de la cible.

Cet article présente les résultats de ces expériences, et une analyse des difficultés techniques rencontrées.

Presented at an AGARD Meeting on 'Virtual Interfaces: Research and Applications', October 1993.

Méthodes

Dispositif expérimental : Le sujet était assis dans la nacelle de la centrifugeuse (figure 1). Sa tête était approximativement placée au centre de l'hémisphère écran (diamètre 1,80m). Sur cet écran, une cible ponctuelle de 1/10° environ pouvait être projetée par une source laser. précis de la tête de façon à placer le centre de rotation de l'oeil droit au centre géométrique de l'hémisphère écran (figure 1). Les mouvements oculaires furent ensuite calibrés en présentant successivement 15 points (5 colonnes X 3 lignes, espacées de 10°). Cette procédure d'initialisation permettait de déterminer l'origine et l'orientation de la LDV issue de l'oeil dans le repère centré par la sphère.

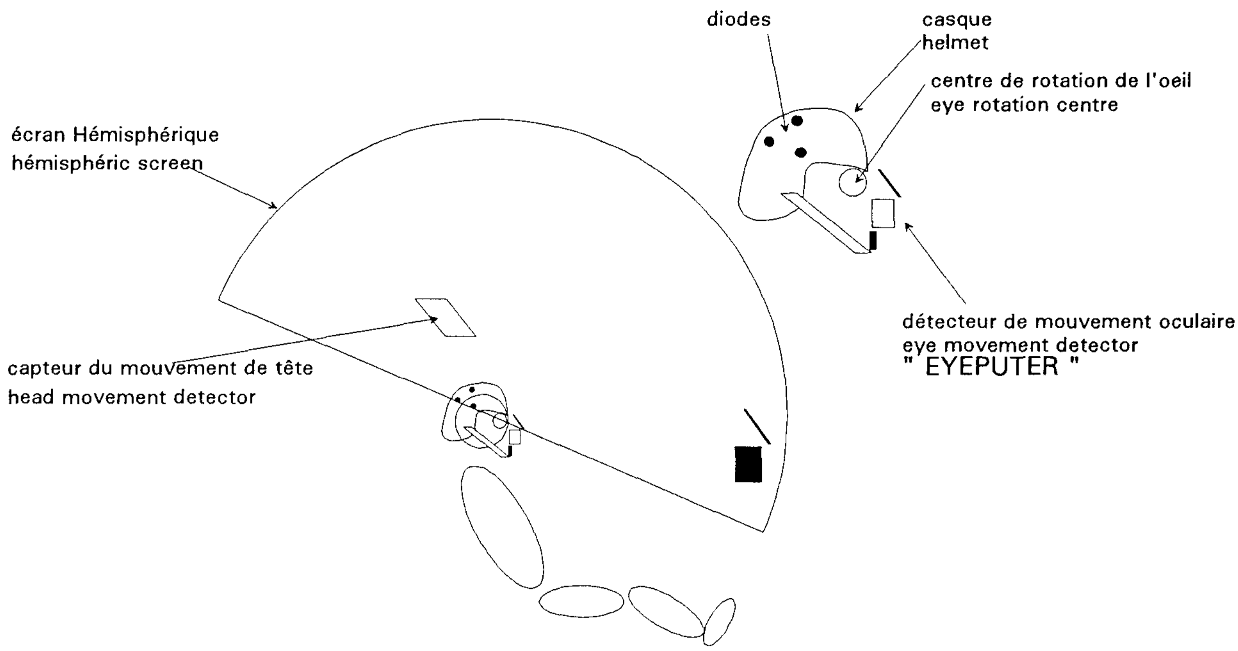


Figure 1: Schéma du dispositif expérimental
et du positionnement du sujet dans la nacelle de la centrifugeuse
Experimental device and subject setting in the centrifuge gondola

Acquisition des données : La LDV était calculée par combinaison des positions Oeil-dans-un-repère-tête et Tête-dans-un-repère-nacelle.

Les mouvements oculaires étaient enregistrés avec un système optique, EYEPUTER, conçu et réalisé par le Laboratoire d'Electronique et de Technique Informatique, filiale du Commisariat à l'Energie Atomique. La précision de la mesure était meilleure que 1°.

Les mouvements de la tête étaient enregistrés avec un système électro-optique, CALVIS, conçu et réalisé par la société Sextant Avionique. Le système délivre 6 degrés de liberté. La précision était meilleure que 1°.

Le sujet portait un casque muni de diodes infrarouges, constituant la partie émettrice du CALVIS, le capteur étant situé au-dessus en arrière du sujet. Par ailleurs, le dispositif EYEPUTER était disposé devant l'oeil droit du sujet, fixé sur une barre reliée au casque et maintenue par rapport au visage par une empreinte dentaire («bite board»). La masse totale du casque, muni de l'oculomètre et des câbles était de 2 kg environ.

Protocole :

La mise en place du sujet débutait par un positionnement

Première expérimentation: La **tâche** assignée aux sujets était de suivre du regard, c'est-à-dire en combinant des mouvements de la tête et des yeux, une cible se

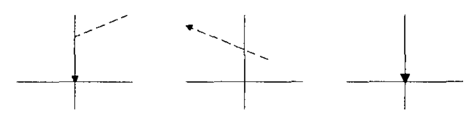


figure 2 : différentes trajectoires de la cible présentée
au cours de la première série expérimentale

déplaçant selon des trajectoires connues. La durée de chaque essai était 10 secondes. Trois **trajectoires** en 2D ont été proposées (figure 2).

Les différentes valeurs d'accélération testées ont été: 1 (nacelle à l'arrêt), 4 et $5G_z$.

Sur le plan matériel, une première version de l'oculomètre EYEPUTER était mise en oeuvre.

Sur le plan logiciel, la détermination de l'intersection de la LDV avec l'écran a nécessité une correction de l'erreur de parallaxe liée à la proximité de l'écran par rapport à la tête.

Seconde expérimentation: Elle différait de la première par le mouvement de la cible et la tâche assignée au sujet.

Mouvement de la cible : La cible se déplaçait sur 8 points choisis sur les axes. Elle restait 1 seconde à sa place puis gagnait la position suivante à faible vitesse (<10°/s).

Tâche : La tâche demandée consistait à suivre du regard la cible en mouvement. Deux modalités ont été proposées aux sujets. Dans l'une le sujet suivait la cible en déplaçant sa tête et ses yeux (OEIL+TETE), dans l'autre, un réticule collimaté à l'infini était positionné devant l'oeil droit du sujet. La consigne était de maintenir ce réticule sur la cible, ce qui l'obligeait à une poursuite tête seule (TETE SEULE). Dans cette dernière condition, le mouvement oculaire n'était pas enregistré. L'oeil restait en position neutre.

Les accélérations testées ont été 1, 1.4, 2 et $3G_z$.

Sur le plan matériel, une seconde version de l'oculomètre EYEPUTER fut utilisée, en association avec un matériel informatique plus performant. Une datation simultanée des données oeil et tête a été utilisée pour synchroniser les signaux.

Sur le plan logiciel, la correction de l'erreur de parallaxe a été modifiée par le traitement entièrement numérique des données, la datation des données et une modification des algorithmes de traitement. Une procédure interactive a confirmé que la précision de la projection de la mesure de la position de la tête sur l'écran est satisfaisante.

Résultats

Première expérimentation :

La figure 3 présente un exemple de la coordination oeil-tête enregistrée dans un essai à 4G mené sur une trajectoire sortante.

Les tracés supérieurs montrent l'enregistrement de la position oculaire en site et en gisement. Ce tracé obtenu avec la première version de l'oculomètre montre une relative stabilité des valeurs. On observe cependant des fluctuations d'allure périodique sur le tracé de site et une dispersion en fin d'enregistrement sur le tracé de gisement, liée vraisemblablement à la désunion des différents éléments (reflets et pupille) utilisés par l'oculomètre. La précision de la mesure observée sur banc permet d'accepter la valeur de la trace principale comme valide.

Les tracés du mouvement de la tête, au centre de la figure, sont caractérisés par la stabilité de la trace et l'absence de fluctuation.

Les tracés du regard, dans la partie inférieure de la figure, sont obtenus par combinaison des traces précédentes. Le déplacement de la cible est représenté par la trace rectiligne. Le regard calculé suit sensiblement, en site, une trace parallèle à la cible. En gisement, le tracé du regard semble plus proche de la valeur cible.

L'aspect performance de l'essai est résumé sur la partie droite de la figure 3. L'erreur de visée en site est retrouvée sur la trace «trajectoire». Elle est confirmée par le tracé de l'erreur instantanée en site et gisement, par rapport à la position de la cible, erreur au-delà de 2° le plus souvent.

Seconde expérimentation :

La figure 4 présente un exemple de poursuite en condition TETE SEULE. Sous une accélération de 2G, le sujet pointe avec sa tête. Le maintient du regard sur la cible est traduit ici par une erreur quadratique moyenne (EQM ou Root Mean Square RMS) de O, 8 en site, 0,9 en gisement et 1,2 d'erreur vectorielle. Le regard n'étant porté ici que par le mouvement de tête, on observe en gisement un dépassement à chacun des arrêts (durée 1s) du mouvement de la cible.

Les figures 5 et 6 présentent un exemple de poursuite en condition OEIL+TETE. L'enregistrement oculométrique, du fait de sa durée (>30s), présente des pics communs aux tracés site et gisement dus aux clignements palpébraux. La figure 6a montre le déplacement du point visé du regard. On constate un décalage constant entre la position calculée du point de regard et la position de la cible. Toutefois la stabilité du regard est assurée malgré les fluctuations compensatoires des positions de l'oeil et la tête. De plus, les débordements observés précédemment à l'instant initial des inflexions du mouvement de la cible sont très amoindris.

Discussion de la méthode et des résultats

La technique consistant à combiner un enregistrement du mouvement oculaire avec un enregistrement du mouvement de la tête est a-priori la seule envisageable dans un environnement aéronautique. La mesure du mouvement de la tête par une méthode électro-optique a été choisie en raison de sa précision et de son insensibilité au milieu, au contraire des méthodes magnétiques difficiles à mettre en oeuvre dans un environnement de type centrifugeuse.

L'intégration des données tête dans un algorithme de correction de l'erreur de parallaxe a pu être validée. La méthode consistait, pour un opérateur muni du casque et du réticule fixé, à maintenir le réticule sur plusieurs points successifs sur l'écran, tandis qu'il effectuait des translations et/ou des rotations de la tête. L'erreur entre la valeur calculée du point regardé sur l'écran et ce point restait très faible. Ce qui est confirmé par les tracés bruts obtenus en condition TETE SEULE, où le sujet doit superposer le réticule sur la cible.

L'oculomètre utilisé présente la particularité d'utiliser plusieurs modes d'éclairage d'un même élément oculaire, pour augmenter la redondance des informations.

La gestion de ces modes d'éclairage présentait dans la première série expérimentale des lacunes qui se traduisait par la fluctuation des valeurs observées. La version actuelle de cet oculomètre a amélioré ce problème. La figure 6a montre cependant la persistance d'un décalage.

Ce décalage variant d'un sujet à l'autre, il apparaît lié à la difficulté pour un sujet de déterminer par lui-même une position zéro de son oeil. Les matériels, utilisés lors des phases d'initialisation de la position de la tête et de calibrage préliminaire des mouvements oculaires, étant incompatibles entre eux, le décalage observé est vraisemblablement dû à une dérive oculaire insensible, entre ces deux phases. La position zéro de l'oeil à l'initialisation de la tête correspondrait alors à quelques degrés lors du calibrage oculaire. Pour vérifier ce point, nous avons procédé à un nouveau calcul de la position de regard après avoir introduit un décalage dans les données oculométriques (figure 6b). Ce rajustement amène les valeurs de l'EQM à 0,8 en site, 0,9 en gisement et 1,2 en vectoriel dans cet essai.

Perspectives

La détection de l'intersection de la ligne de visée avec les objets de l'environnement est une perspective séduisante qui pose encore de nombreux problèmes. Les capteurs utilisés ici devraient bénéficier d'avancées techniques, en terme de fiabilité et stabilité des valeurs obtenues, de fréquence de mesure et de diminution de la masse supportée par la tête. L'intégration de ces capteurs dans une boucle de commande pose le problème du traitement des données (par exemple le rejet des artefacts). Il se pose le problème de l'ergonomie de ce type de dispositif. Les expérimentations décrites ici ont demandé une procédure d'installation extrêmement rigoureuse, liée à l'hétérogénéité des capteurs employés et leur complexité.

En dernier lieu, la stabilité de l'oculomètre face à l'oeil nécessitait l'emploi d'une empreinte dentaire.

Si une telle méthode est possible en laboratoire, en pratique, les dispositifs envisagés devront prendre en compte leurs procédures d'emploi dès la conception. Compte-tenu des installations déjà implantées sur les casques, un système porté sur la tête devrait constituer une entité dont la réalisation relève de l'industrie, et la validation du laboratoire.

Conclusion

Dans cette expérimentation pilote, l'intersection de la ligne de visée d'un sujet avec un écran hémisphérique le recouvrant a été calculée en centrifugeuse. La faisabilité de l'expérience et les résultats préliminaires obtenus montrent l'intérêt de la méthode, comparée à la désignation par la direction de la tête. Un développement ultérieur pourrait bénéficier des enseignements futurs, ergonomiques et physiologiques, que l'usage d'un tel dispositif permet d'espérer.

Remerciements: Les auteurs remercient Monsieur Dominique MASSE et son équipe du CEA/LETI pour leurs conseils et Monsieur René VALLET du CEV-LAMAS pour son très important travail technique.

Ce travail a été réalisé grâce à la commande 89/1528 de la Direction des Recherches Etudes et Techniques .

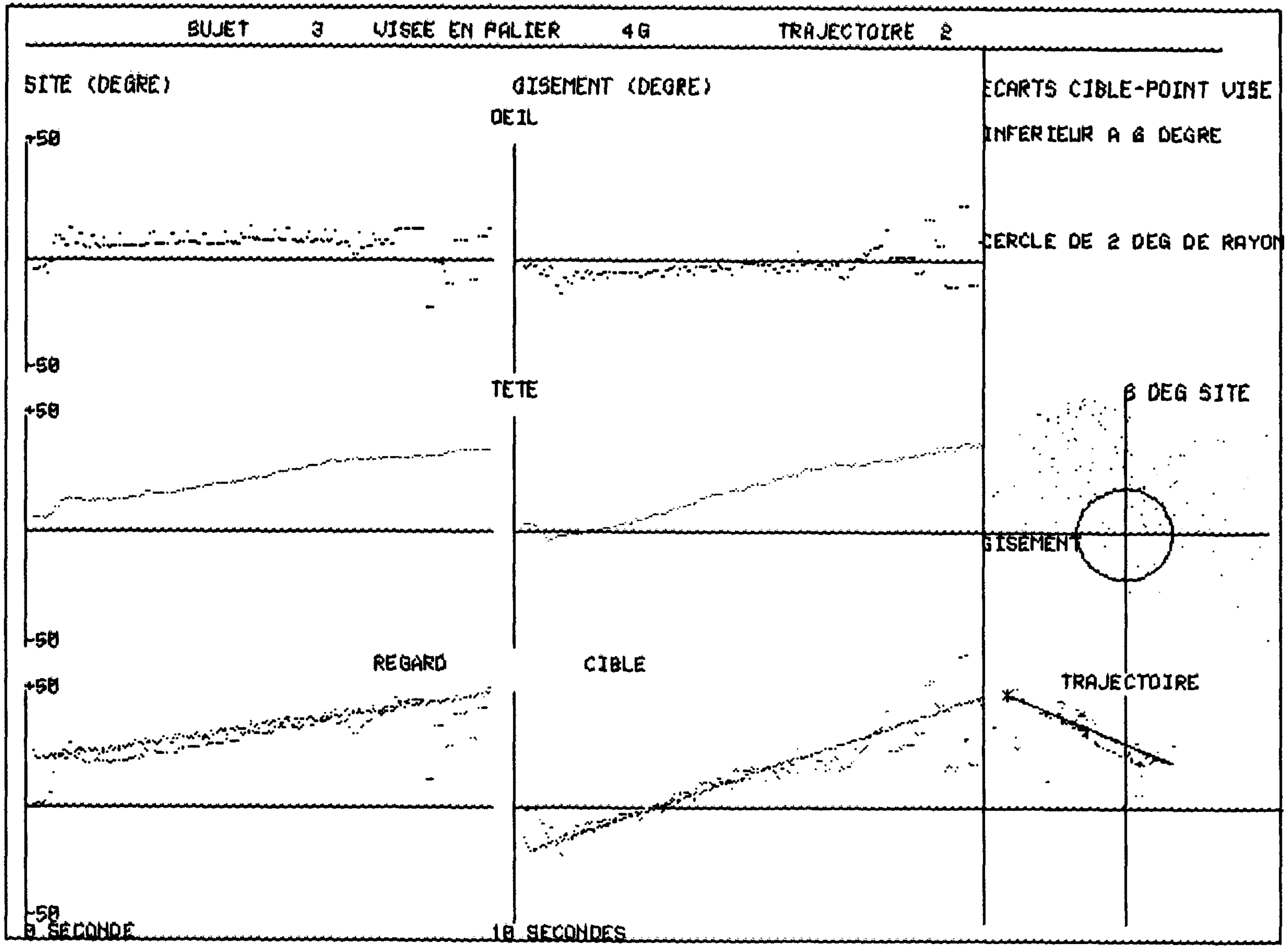


Figure 3 : OEIL+TETE, 4G, trajectoire sortante, exemple de coordination Oeil-tête
EYE+HEAD, 4G, exiting trajectory, Eye-head coordination example

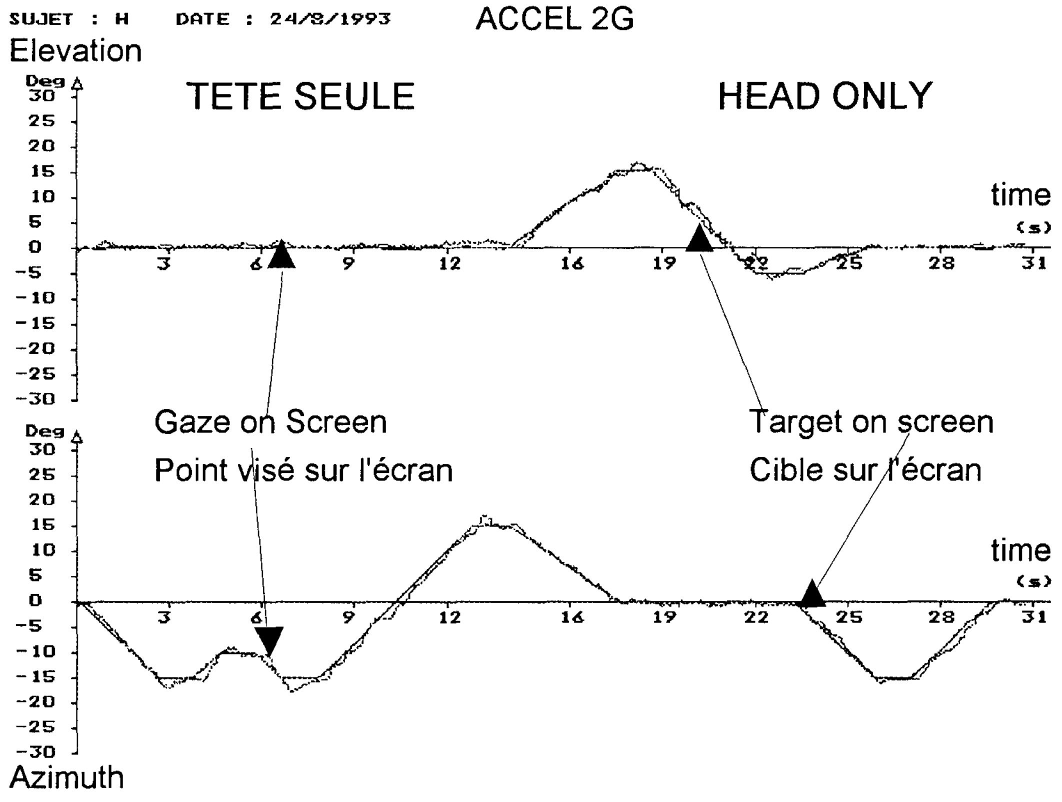


Figure 4 : Tête seule, 2Gz, point visé sur l'écran en fonction du temps
Head Only, 2Gz, point of gaze displacement on the screen vs time

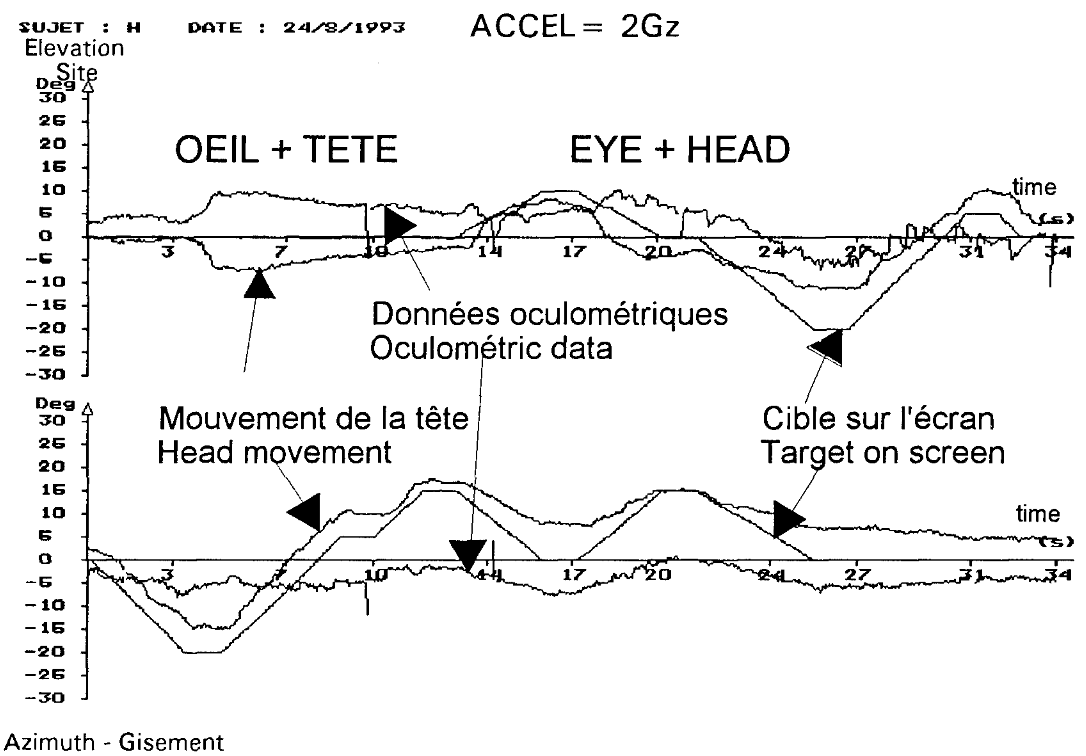


Figure 5 : Oeil+Tête, 2Gz, tracés des déplacements oeil et tête en fonction du temps
Eye+Head, 2Gz, eye and head movements vs time

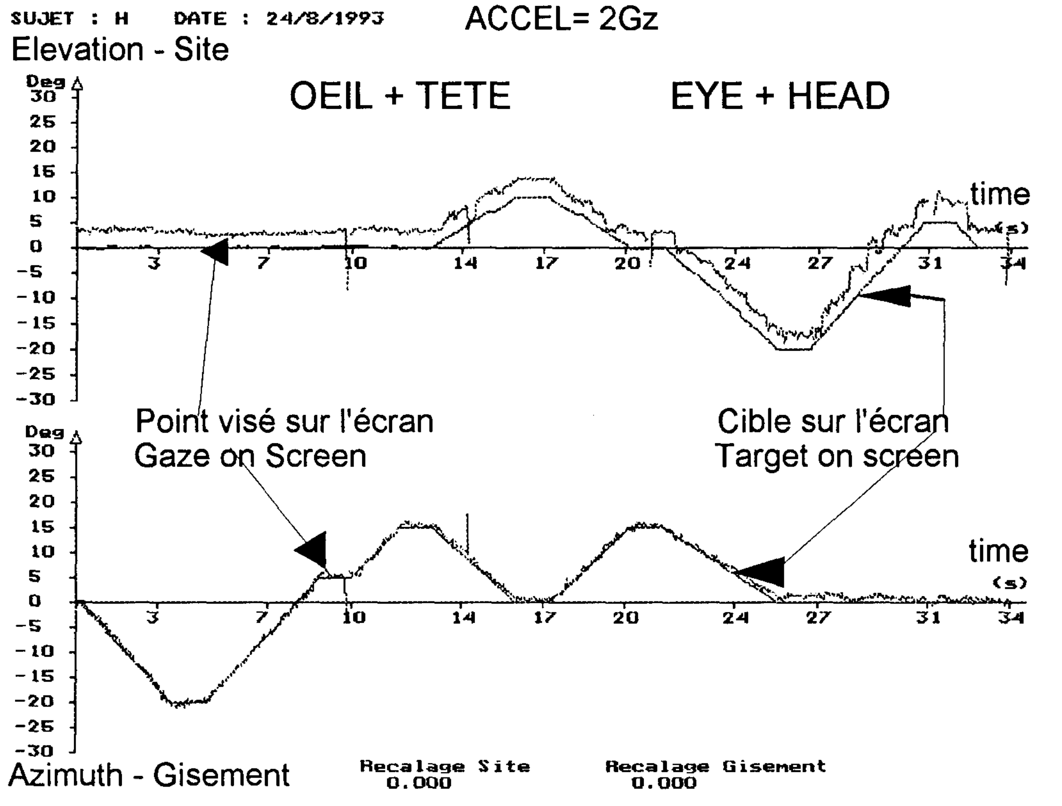


figure 6a : Oeil+Tête, 2Gz, déplacement du point visé sur l'écran en fonction du temps données brutes, les traces sont parallèles, non superposées.

Eye+Head, 2Gz, point of gaze displacement on the screen vs time raw data; note that traces are parallel, not superimposed.

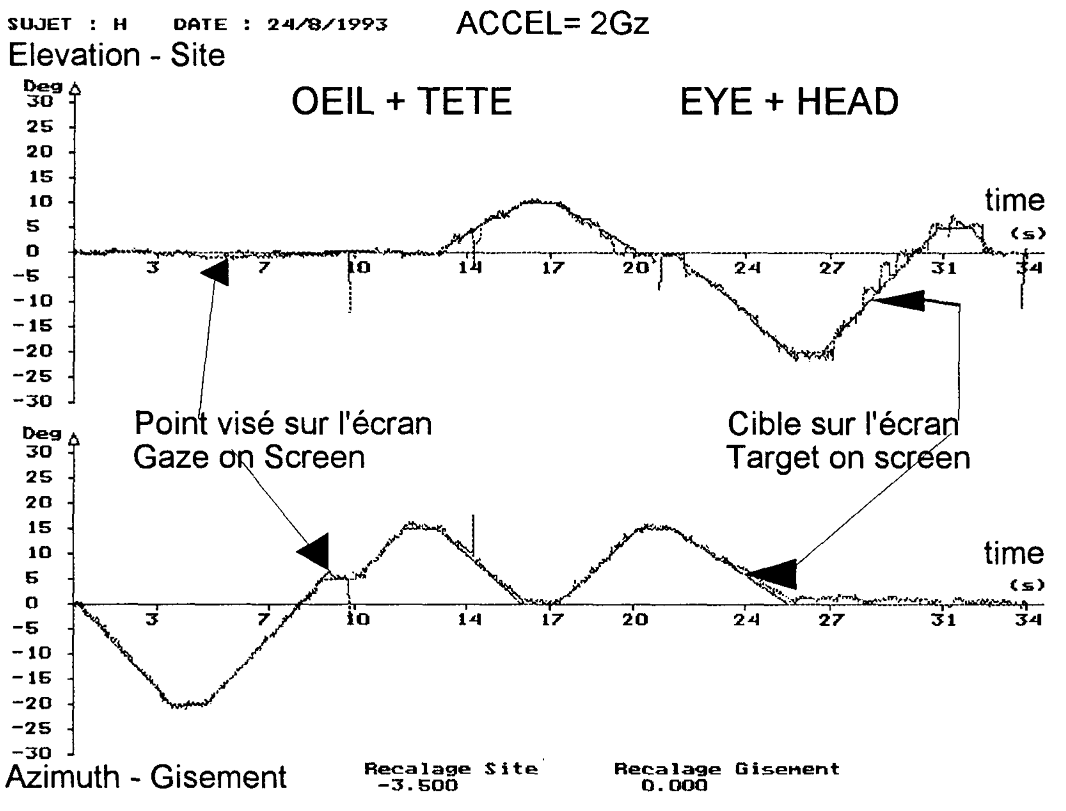


figure 6b: Oeil+Tête, 2Gz, déplacement du point visé sur l'écran en fonction du temps
Après recalage (ici -3.5°) les traces sont superposées.
Eye+Head, 2Gz, point of gaze displacement on the screen vs time
When shifted (here -3.5°) traces are superimposed.

A Comparison of Two Examples of Magnetic Tracker Systems

M. Williams

Human Factors Department
Sowerby Research Centre
FPC 267 British Aerospace
Bristol BS12 7QW
United Kingdom

SUMMARY

This paper is an account of an investigation of the performance of various position measuring devices which use low frequency AC or pulsed DC magnetic fields. They are used in many applications in computer graphics, and now for "Virtual Reality", where it is necessary to estimate the observer's direction of gaze. As part of the Sowerby Research Centre's programme of eye movement research one such system is being integrated with a video based eye-tracker.

There seems to be no independent, published assessment covering all aspects of all the systems which are of interest to this research programme. This paper aims to fill that gap: it includes information relating to the static performance of two measuring systems: the 3-Space Polhemus Tracker and the Ascension Technologies' "Bird". The measurements relate to repeatability, noise, cross-talk, stability, range and linearity. The influence of metal objects close to the transducers is also investigated. In most respects the "Bird" sensor was found to be more appropriate for this application.

INTRODUCTION

Since the early experiments of Ivor Sutherland, (Ref. 15), using mechanical linkages and subsequently ultrasonic range-finders, there have been many attempts at making devices which can inform a computer of the locations and orientations of its user's limbs. The importance of the fidelity of such information increases as the demand arises not just in Interactive Virtual Environments (IVEs) but also in weapons aiming systems (examples of the Polhemus magnetic tracking system, a commercial version of which is investigated here, were flown in F4s in 1972 see also Ref. 13) In an attempt to develop an environment more suited to a pilot's tasks in the future, means for providing a completely enclosed cockpit have been sought (for example the "Super-Cockpit" Programme at Wright-Patterson Airbase and also the VCASS or Visually Coupled Airborne Systems Simulator). This usually implies that a pilot will fly using cues which are synthetic and provided by computer generated imagery. In an alternative scenario it may be necessary to provide computer enhanced or highlighted imagery from, for example, infra-red sensors. In both of these cases it may be necessary to determine where the pilot is looking (e.g. Ref. 17) and any inadequacies in the tracking systems may be apparent as mis-registration of imagery or as image lag.

Magnetic tracking systems coupled with Helmet Mounted Displays (HMDs) have been reported by many Virtual Reality researchers from low cost aids for disabled children (Ref. 12) to the NASA Ames HMD project used for tele-robotics applications (Ref. 8). A less "conventional" use of such trackers was, in combination with VPL's Dataglove, to drive a speech synthesiser via a neural network gesture recogniser (Ref. 7). The VCASS system mentioned above had tracking systems specially manufactured for it (Ref. 14).

At British Aerospace's Sowerby Research Centre work is being pursued relating to eye movements and direction of regard in free-space, particularly in relation to eye-pointing tasks. Ideally an accuracy of better than 0.1 degrees and a few millimetres is required. A video based eye tracker is used which is helmet mounted

Presented at an AGARD Meeting on 'Virtual Interfaces: Research and Applications', October 1993.

and so, in order to ascertain direction of regard in free space, it necessary to know the position and orientation of the helmet wearer's head. The problem is more subtle than this since it has been shown that significant amounts of helmet slippage can be expected to occur during voluntary head movements (Ref. 11). In our circumstances it is necessary, therefore, to also ascertain the orientation of the helmet with regard to the wearer's head and so two, simultaneous, measurements need to be made. For some tracking systems this means sacrificing measurement rate because multiple sensors are multiplexed and so it has been necessary to find more suitable ones.

Although comprehensive summary surveys of tracking technologies exist (Ref. 10), with the exception of Adelstein et alia, there has been little detailed work published by independent assessors of these tracking systems appropriate for the work in BAe's Research Programme. Partial investigations are documented in several papers (Ref. 4, 5, 6, 9 and 16). The present paper describes work carried out to contrast the static precision and accuracy of the data generated from two commercially available magnetic tracking systems, it complements work presented in reference 1, which deals with some of the dynamic properties of both of these trackers.

The Polhemus 3–Space

Until recently the only commercially available magnetic tracking system has been the "Polhemus 3-Space" tracker. This is a flexible system which has options in firmware and hardware which allows the tracking of six degrees of freedom and also allows the device to be configured as a "digitiser". In this context, the latter means that a pointer may be run over the surface of a body to allow its geometry to be recorded as a series of vertices and perhaps incorporated in a computer model. The Polhemus consists of an electronics unit and, depending on the model, may have one transmitter and receiver connected to it or up to two transmitters and four receivers. The receivers have dimensions 25 x 15 x 10 mm and the transmitter has dimensions approximately 65 x 35 x 35mm. The electronics unit communicates with a host computer across either an RS-232C serial link or across a proprietary eight bit parallel bus.

The data returned from the device are in the form either of simple x,y and z co–ordinates in space plus orientation angles (roll, azimuth and elevation), or as quaternions which encode the orientation as a vector in space. The Polhemus transmits its data either in ASCII form, as literal string representations of the numbers with five significant figures (a total of forty five bytes per record for six degrees of freedom – x, y, z, azimuth, roll and pitch), or in "binary format". The latter is an encoded format where the data is transmitted as a guard byte with its most significant bit set on, denoting the start of a new record, and then the data for each degree of freedom is transmitted with their most significant bits set to zero. The "missing" bits are placed in the guard byte. This requires only fifteen bytes per record. With a maximum baud rate of 19.2kbaud, it is possible to achieve a 60Hz sampling frequency of all six degrees of freedom only in binary mode. A significant drawback for this system in our application is that using multiple sensors causes the effective sampling rate to be reduced because of the previously mentioned multiplexing effect (this drawback is still present when using multiple electronics units because the transmitter fields have to be synchronised).

The following two tables, (1 & 2) describe the manufacturer's claimed properties of the 3–Space Isotrak and the 3–Space DigitiserTracker.

Recently a competing system has appeared, manufactured by Ascension Technology, called "The Bird". Not to be outflown, Polhemus have replied with another system whose details were not known at the time of writing but pre–launch specification claims appeared in the press of 100Hz sampling rate and "improved accuracy and range".

The Bird – an extended avian simile

The Bird has been designed with Virtual Reality (VR) applications in mind. Its technology is based on pulsed DC magnetic fields. The purpose of this is to reduce the effects of eddy currents in metal objects proximal to its receiver and transmitter (as the magnetic field is held constant during the D.C. part of the "D.C. pulse", the eddy currents die away exponentially). Details of the design of the system and how it subtracts out the effects of the Earth's magnetic field are described in U.S. patents 4,849,692,(Ref. 2) and 4,945,305 (Ref. 3).

The Bird cleverly uses field coils in pairs (e.g. X and Y, Y and Z etc.) in order to increase the signal strength at the measuring position without increasing the current demands of the generating antennae or the necessary flux density in each coil. The measurements the Bird makes may be synchronised to a CRT synch signal. In this mode it may sample at up to 144Hz and is configured as a single transmitter (whose dimensions are 80 x 80 x 80mm and is thus significantly larger than the equivalent device for the Polhemus system) and a receiver, both attached to a system box. The system gets its power from a remote DC power supply, perhaps to reduce mains-borne or re-radiated interference from the transformer.

The system box contains an 80186 processor, plus ancillary communications hardware which supports RS–232 and RS–485 protocols. The latter is also used for inter–communication between multiple "Birds" when it is known as a "Fast Bird Bus". Communication on the serial bus at up to 115.2kbaud is possible with the

Table 1: Manufacturer's specifications for Polhemus 3–Space Isotrak

Parameter	Range (Radial Distance)	Values
Static Angular Accuracy	0– 30 inches	0.85° RMS
Static Positional Accuracy	4 – 15 inches	0.13 inches RMS
	15 – 30 inches	linear degradation to 0.25 inches RMS at 30 inches
Resolution (Angular)	0 – 30 inches	0.35° RMS
Resolution (Positional)	0 – 15 inches	0.09 inches RMS
	15– 30 inches	degrades linearly to 0.18 inches RMS

Table 2: Manufacturer's specifications for Polhemus 3–Space Digitiser / Tracker

Parameter	Range	Values
Static Angular Accuracy	+/- 4 – +/-28 inches	0.5° RMS
Static Positional Accuracy	ditto	0.1 inches RMS
Resolution (Angular)		0.1° RMS
Resolution (Positional)		0.03 inches RMS

option of using a user supplied clock signal for "odd" baud rates.

When multiple units are linked together the system is known as a "Flock of Birds" and they are intelligent enough to enable control to be automatically passed from one system to another as the receiver moves about amongst multiple transmitters. In addition the transmitter power is varied depending on the range of the closest receiver. This means that the proximity of one receiver may affect the noise levels in other receivers in the "Flock". The manufacturer states that maximum power occurs at ranges beyond 9.5 inches and so it is desirable to keep both receivers at greater distances than this figure. Closer than this the power is halved at predefined ranges. Transmitters of higher power are offered to give individual ranges of two metres when the system is known as a "Big Bird"[1]. Table 3 describes the specifications given by Ascension for their system.

EXPERIMENTAL APPARATUS

The apparatus for this investigation consisted of a Polhemus 3-Space Tracker system, a single system unit from a Flock of Birds and a host PC (Viglen VIG IV). In this case no use of Polhemus's "magnetic environment mapping" service. The measuring systems were mounted in turn on a plastic cradle in which a laser pointer could be installed. The whole of the magnetically sensitive system was mounted on a large wooden sheet supported 15cm from the floor by wooden blocks and a cross whose two arms were 60 inches long was inscribed on the sheet's surface. Reference marks were placed at three inch intervals along each arm and to a precision of 1/32nd inch.

The cradle consisted of two rectangular perspex sheets: one to act as a base with a pivot at its centre and with a leveling screw at each corner; the other with three brackets whose internal diameters were the same as the laser pointer. The brackets were aligned so as to support the laser and the middle one was co–incident with the pivot and centre of rotation. This middle bracket had a top section which bridged the laser and onto which a perspex block was fixed, drilled so that the centre of the Bird receiver coincided with the pivot. Three sets of fixing holes were drilled so that each axis of rotation could be measured.

Since the laser used as a pointer in these experiments had an aluminium body the middle bracket had spacers between its lower half and the bridge above. This allowed the laser to be lifted up and passed through without disturbing the rest of the cradle. In addition, the spacers allowed sample metal plates to be inserted in the middle bracket parallel to the receiver in order to get an indication of the field distortions induced by consequent eddy currents in the plate. Two types of plate were considered: a solid aluminium plate and a plate made from three laminations of aluminium approximately the same total mass and dimensions, (approximately 85mm square sides and 3mm thick).

[1] The extended "avian simile" becomes progressively more tedious as references arise to "beware of knocking the birds off their perches" if a certain mode of fixing is adopted!

Table 3: Manufacturer's specifications for the Ascension Technologies' "The Bird"

Parameter	Range	Values
Angular Accuracy	+– 36 inches	0.5° RMS
Positional Accuracy	+– 36 inches	0.1 inches RMS
	15 – 30 inches	linear degradation to 0.25 inches RMS at 30 inches
Resolution (Angular)	at 12 inches	0.1° RMS
Resolution (Positional)	at 12 inches	0.03 inches RMS

EXPERIMENTAL METHODS

General

In all the measurements described below, the "physical x–axis" was aligned parallel to the x–axis of the transmitter (as defined in the respective user manuals) and the transmitter and receiver oriented so that a displacement of the receiver upwards introduced a positive z measurement. The relative height of the transmitter was adjusted until the receiver indicated that it was displaced only along the x–axis, ie. all other measurements were indicated as zero or as close as could be obtained in the case of orientation measurements. The *transmitter's* position was then left unchanged throughout all subsequent measurements. In the case of the Bird, it was found that the centres of measurement were best interpreted as being the geometrical centre of the transmitter and the mid point of the *left hand side* of the receiver (the receiver being viewed from above with the cable exiting towards the viewer and its mounting brackets being towards the bottom).

The general method was to displace the receiver along one axis at a time. When measuring the x–axis, for example, an initial displacement of nine inches was used. The receiver's position was recorded and then it was displaced a further three inches and a measurement recorded again. This process was repeated at three inch intervals until the receiver had been displaced a total of thirty nine inches, the last measure was repeated and then the receiver was moved back towards its original position in three inch steps with data being recorded at each point as before.

The exceptions to this procedure were investigation of the variation with time (where the receiver was simply left in its initial position for an hour) and the "metal proximity" measurement (where the receiver and transmitter were left in their initial positions and a metal plate moved between the two). Table 4 summarises the measurements made. Some of these measures were repeated after a week but were found to be consistent.

When measuring with the Bird, each "measurement" consisted of ten values taken at the default sampling rate of 100 Hz and from which a mean and standard deviation were formed. The latter was used to give some measure of the noise in the signal. With the Polhemus, the same approach was taken except that binary mode was used so that the effective sample rate was 60Hz. For both systems, the data were returned with the unit in continuous transmit mode, (rather than requesting each individual measurement).

Prior to each experimental "run", which would consist, for example, of measurements with a metal plate in place bracketed by two sets of measurements without metal plates, the Polhemus system was boresighted to define the orientation axes to be those of the current physical orientation of the receiver.

The whole of the experiment was controlled from software written by the author using Microsoft 'C' version 7.0.

Units

In all cases the natural units of measure, for these systems, were used i.e. translations were measured in inches, rotations were measured in degrees.

Stability in time

This consisted of simply arranging for the experimental software to take a measurement every thirty seconds over the period of an hour. Timing was performed using the PC system clock.

Repeatability

In each case a single co–ordinate was increased and subsequently decreased while measurements of all six degrees of freedom were noted. Each measurement was therefore made twice at each calibration point on the cross.

Filter effects

Essentially similar to the repeatability measures, the effect of "turning off" various software filters in the Bird system unit was investigated using the repeated measures given above.

Eddy current effects

Two aspects were investigated: firstly what we term "metal proximity". Small samples of aluminium were used (approximately 85 x 85 x 3 mm) in various ways. Measurements were made with the transmitter and receiver displaced only along the x–axis and the sheet of metal was placed between the two, flat on the measuring table lying symmetrically about that axis. The distance of the metal plate was varied and this series repeated with a second separation of transmitter and receiver. In addition measurements were made with the metal sheet standing on end, perpendicular to the X–axis

The second aspect of eddy current effects investigated is referred to here as "field distortion". As with the preceding set of measurements, this was a difficult thing to measure in some systematic and meaningful way: in this instance, the metal plate was placed *in the cradle* used to support the receiver, between the receiver and transmitter with the plate's surface parallel to the table. X and y were varied separately and the effect on all variables examined. Several combinations were tried, measuring displacement along the x and y axes separately: these were filters on or off; metal plate present or not; and finally with the metal plate aligned with the normal to its surface pointing in the X direction. In the case of the Bird additional conditions were used to examine the effects of the software filters available.

RESULTS AND ANALYSIS

The results from these measurements were rendered graphically using the UNIRAS graphing system for a Sun platform. This package allows limited statistical analyses, which seem adequate in this instance.

Variation with time

The graphs presented here are representative of measurements made. In the context of static measurement over a period of an hour, no significant variation in the measured values of co–ordinates was observed with the Bird for the default values for filters A few "blips" were observed in the Y variation, these however were very small – the difference measured was 0.01 inches and so were below the claimed resolution of the system for this range. The same was true for the Polhemus, although minor drift seemed to manifest itself in the last twenty minutes of measurement, (Figures 1 & 2). This difference may be a feature of different internal filtration methods.

The various repeatability measurements made (for example the metal / non-metal measurements separated by at least a week, or those examining the effect of the presence of the filter when a metal plate is close to the receiver) indicated that the calibration remained good with the Bird (Figures 9 – 14) and (Figures 15 – 20). The same seems to hold true for the Polhemus, however assessment of the long term stability has not been made for an interval greater than a day.

Metal Proximity – Bird only

Two conditions were tested with the receiver – transmitter separation of fifteen and twenty inches. As with other conditions, angular measures were most sensitive to any effects shown, the distortions produced were most obvious with increased separation of transmitter and receiver. In all cases the apparent displacement was less than 0.1 inches in all axes. Orientation measures were disturbed to a greater extent with deflections of up to 0.4^{o} (Figures 3 – 8). If the receiver was unperturbed by the presence of the metal plate, one would expect no variation in the measured values.

Field Distortion – The Bird

Figures 9 to 14 show the variation in the difference between a variable's measured values for the conditions with a plate and those measured without; plotted as a function of the range of the transmitter. The measurements were repeated a week later and are indicated by the subscript "1". Ideally, one would expect these plots to be straight lines of slope zero. The presence of the metal plates affected mostly the measure in the Z direction, i.e. perpendicular to the surface of the plate, and the pitch measure.

Figures 15 – 20 demonstrate the presence of various metal plates: subscripts are "l" for laminate, "m" solid metal plate and "p" for the metal plate perpendicular to the X axis. Again one would expect these graphs to be linear with zero slope. Insertion of a laminated rather than a solid metal plate had the effect of reducing some distortions– but not greatly. This effect would only be significant if the laminate was formed so that non–conductive layers formed the greater part of the construction.

(Figures 21 – 26). These graphs illustrate the effects of the filtration on the measurements and the "quasi–periodic" nature of the error. There seems to be no or negligible effect on the absolute value measured, however, as discussed below, there is an effect on the noise generated. In these cases, filters off means that all filters were switched off: DC, Narrow band AC and Wide Band AC.

Field Distortion – The Polhemus

The Polhemus exhibited similar effects to the Bird and errors induced by the proximity of metal seemed to be of the same order if marginally larger than the Bird, which is surprising since it had been claimed that the Bird would be an improvement. The results seemed swamped, however by what appear to be larger non–linearities inherent in the measurement system, (Figures 27 – 32).Compare for example Figure 9 with Figure 27. The Bird exhibits about half the difference in

Table 4: A Summary of Measurements made
"X" indicates measurement was made

Parameter	Polhemus	Bird
X linearity and repeatability	X	X
Y linearity and repeatability	X	X
Across parameter "cross-talk"	X	X
Metal Proximity	–	X
Field Distortion: solid metal in X–Y plane	X	X
Field Distortion: solid metal in Y– Z plane	X	X
Field Distortion: laminate in X–Y plane	X	X
Signal Noise: effect of software filters	–	X
Signal Noise: effect of metal proximity	–	X
Signal Noise and Range	X	X
Time stability of Signal Noise and accuracy	X	X

error in X that the Polhemus does and does so in a linear fashion.

The distortions introduced by the metal plate are most obvious in the Polhemus in the pitch measurements (when using x as the independent variable) where deviations of several degrees are observed at the extremes of the measured ranges. In other cases the individual spread of the trajectories is too great to allow definite isolation of the effects of the metal plate. Other differences were apparent when comparing the effect on all variables of varying Y (Figures 33 – 38). In this case the Polhemus seemed to behave less well, its orientation measurements varying as much as the *unfiltered* values from the Bird. The poorest performance seems to be in measuring roll, where the amplitude of the non–linearities is approximately 4° alone! (Figure 37)

Signal Noise

The results for the Bird are summarised in the Table 5. For static measurements, as presented here, small but measurable differences in the amount of noise present could be discerned when metal was introduced close to the receiver and the filters were turned off. With filters switched on the difference in noise levels were almost unmeasurable whether metal was present or not, turning off the filters produced an increase in noise which was a function of range. When metal *was* present, however, the mean increase in noise (ie averaged over the domain of separations) was approximately twice that of the receiver without metal present, (Table 5).

The Polhemus system was tested only with filters present: in this case, as with the Bird, there was little difference between the two conditions where metal was close to the receiver and with no metal close.

Range

With both systems there is a trade off of noise against range. With standard configurations, however, it is clear that the Bird has a "hard limit" of around 36 inches whereas the Polhemus exhibits a gradual deterioration to 60 inches. It has not been tested to 60 inches but this is the manufacturer's claim. The evidence in the measurements which have been made certainly supports the assertion that noise increases with range, the apparent *error* in X did not increase linearly, however.

DISCUSSION

The results presented here are consistent with those presented elsewhere for a Polhemus Isotrak (Ref. 5). Burdea et al. (Ref. 5) did not test their system to the same range as presented here but those data which can be compared are in agreement. In addition, there is some evidence for errors introduced to measured "X" displacements by physical displacements in "Y" as reported here but, perhaps because the measurements were carried out at a greater "X" range then those reported by Burdea et al., the effects themselves are greater. Bryson (Ref. 4), also presents data which

Table 5: Comparison of the increases in means of S.D. of noise when filters are turned off with and without metal plate (for Bird)

Parameter	Without metal	With Metal
X	0.07	0.14
Y	0.04	0.08
Z	0.04	0.08
Yaw	0.17	0.30
Roll	0.13	0.22
Pitch	0.16	0.28

is consistent with results given here however the author does not include measurements for orientation nor does the author give any data relating to the Ascension system.

With the Bird, as was to be expected, the presence of filtering has a substantial effect on the noise in the data but this might have had some implications for induced lags. In fact Adelstein et al.(Ref. 1) have shown that the presence of the default filters does not affect latency rather the reporting of orientation with position as opposed to just position was far more significant. The best latency performance, in the examples they tested, seemed to be given by an early Polhemus Tracker. This particular model had a custom EEPROM with all internal filtering eliminated. For units comparable to those tested here, the Ascension system appeared to have inferior latency properties for tracking low frequency stimuli (less than 2.5Hz) but this response was held flat for all frequencies whereas the Isotrak frequency dependency was more complex.

Aside from the lags, it would seem that there is a possibility of introducing corrections into the measuring system if metal is excluded from the vicinity of the transmitter and receiver since most deviations of the measurements seem to be nearly linear or quadratic. Clearly, the results pertaining to the effects of the proximity of metal are only an *indication* of what can be expected, since, as has been shown, the orientation of the plate with respect to the receiver and transmitter can have some effect on the measurements. However, provided the increase in noise is tolerable and provided that any conductive structure and measuring system have a fixed geometrical relation a relatively simple second order correction may suffice even here.

SUMMARY OF RESULTS

- Both magnetic systems seem stable over time.
- Angular measurements are most sensitive to any non-linearities either in the system or induced by the proximity of conductive objects
- Errors induced were sensitive to the orientation of metal plates. Typically less than 0.4 inches at maximum range and angular errors of +/- 0.5 degrees
- Unlike the Polhemus, the Bird has a "hard limit" maximum range of about 36 inches
- The centre of measurement for displacement seems to be the mid point of one side of the receiver for the Bird but the geometrical centres of transmitter and receiver for the Polhemus.
- For the Bird, switching on its software filters had negligible effect on the errors induced by the presence of metal plates, it did however reduce the noise induced by approximately one half.
- Larger "cross–talk" effects were observed with the Polhemus, particularly with respect to orientation measurements.

Description of Graphs

The following graphs illustrate the measurements made on the Bird and Polhemus. Generally, dependent variables are labelled with names which indicate the special condition that they represent. For a fuller discussion, please refer to the "Results" section of the paper. For both the Polhemus and Bird the measurements broadly fall into three groups. These groups are: simple time variation of measured co–ordinates with a static receiver and transmitter (Figures 1 & 2 for Polhemus only), the variation of all six degrees of freedom as the receiver is displaced along the x–axis and lastly the effects on all six–degrees of freedom when the receiver is displaced along the y–axis.

These groups may be further subdivided. Figures 3 to 8 have the apparent position (i.e. values returned by measuring device) of the receiver plotted as a function of the separation of metal plate from the receiver along the X – axis. Figures 9–14 plot the differences between the measured values with and without the presence of a metal plate on two separate occasions (separated by a week).

Figures 15 to 20 show the measured values as a function of the x–displacement of the receiver when the type or orientation of the interposed plate is changed. In the case of the X measure, the "real" value has been subtracted to allow the actual variation to be plotted clearly. Figures 21 to 26 illustrate the effect of the Bird software filters and metal plates this time as a function of y – displacement of the receiver.

Table 6: Index to graphs

Figures	Parameters
1,2	Examples of Time Variation in Polhemus Measures
	Bird
3–8	Changes in measured co-ordinates with movement of metal plate along the x axis between the receiver and transmitter. Ordinate is displacement of metal plate from transmitter
9–14	Variation in measured co-ordinates with x-displacement of receiver when metal plate was mounted parallel to X–Y plane in cradle (Two measurements per graph of same measured co-ordinate)
15–20	Effects on measured co-ordinates with x-displacement of receiver with various plate configurations (Includes: no metal, solid aluminium in X–Y plane, laminate in X–Y plane and solid plate in Y–Z plane)
21–26	Changes in measured co-ordinates with displacement in Y axis of receiver (effect of metal plate in X–Y plane and filters included)
	Polhemus
27–32	Effects on measured co-ordinates with X displacement and metal plate in X–Y plane (as for Bird figures 9–14 and 15–20)
33–38	Effects on measured co-ordinates of different material types for comparison with figures 21 –26 (Solid aluminium and a laminate included)

For the Polhemus, because preliminary measurements showed a degree of variation in the measured results, several measurements have been bracketed around the repeated measurement of the effect of the metal plate. In all there are five measurements of returned values as a function of x and two with returned values showing the effects of metal plates. There are two groups of measurements separated by a day, (Figures 27–38).

Graph Legends

The following abbreviations are used in the labelling:

diff The value plotted is the difference between the sensor's measured value and the "correct" physical value.

1,2,3... Usually these represent repeated measurements of the same case

1&2 In the case of measuring "metal proximity" these sub-scripts represent two separations of receiver, the graphs then illustrate the *difference* of returned values from some arbitrary base.

m A case where a solid aluminium plate was used in mounting cradle with plate parallel to X–Y plane

l A case where a laminated aluminium plate was used as with item above

p A solid metal plate was used but the plate was orientated perpendicularly to the table and x–axis – i.e. parallel to Y–Z plane

f & m Used when measurements were made with metal close to the receiver f – filters in place, m with no filters

References

1. Bernard D. Adelstein, Eric R. Johnston and Stephen Ellis, 1992, "A Testbed for Characterizing Dynamic Response of Virtual Environment Spatial Sensors", 5th Annual ACM Symposium on User Interface Software and Technology, Monterey,California, ACM 15–22.

2. Ernest B Blood, 1989, "Device for quantitatively measuring the relative position and accuracy and orientation of two bodies in the presence of metals utilizing direct current magnetic fields", U.S. Patent 4,849,692

3. Ernest B. Blood, 1990, "Device for quantitatively measuring the relative position and accuracy and orientation of two bodies in the presence of metals utilizing direct current magnetic fields", U.S. Patent 4,945,305

4. S. Bryson, 1992, "Measurement and Calibration of Static Distortion of Position Data from 3D Trackers", SPIE, Vol. 1669, Stereoscopic Displays and Applications III.

5. G.C. Burdea, S.M. Dunn, C.H. Immendorf and M. Mallik, 1991, "Real Time Sensing of Tooth Position for Dental Digital Subtraction Radiography", IEEE Trans. Biomed. Eng., BME-38 (4), 366–378

6. J.C. Chung, M.R. Haris, F.P. Brooks, H. Fuchs, M.T. Kelley,J.Hughes, M.Ouh–young, C. Cheung, R.L. Holloway and M. Pique, "Exploring Virtual Worlds with Head mounted Displays",1989, SPIE Vol. 1083, Three Dimensional Visualisation and Display Technologies.

7. S. Sidney Fels and Geoffrey E. Hinton, 1990, "Building Adaptive Interfaces with Neural Networks: The Glove–Talk Pilot Strategy", 1990, in Human Computer Interaction Interact '90, (eds. D. Diaper et al.), p.683-688, Elsevier Science Publishers

8. S.S. Fisher, M. McGreevy, J. Humphries and W. Robinett,1986, "Virtual Environment Display System", Proc. 1986 Workshop on Interactive 3D Graphics, p.77–87

9. Jiandong Liang, Chris Shaw and Mark Green,1991, "On Temporal–Spatial Realism in Virtual Reality Environment", ACM UIST '91,

10. Kenneth Meyer, Hugh L. Applewhite and Frank A. Biocca,1992, "A Survey of Position Trackers", Presence, Vol 1, No. 2, Spring 1992

11. C.N. Neary, I.J. Bate, L.F. Heller and M. Williams, "Helmet Slippage During Visual Tracking: The Effect of Voluntary Head Movement", Aviation, Space and Environmental Medicine, July 1993, p623–630

12. Randy Pausch, April 1991, "Virtual Reality on Five Dollars a Day", ACM SIGCHI Conference Proceedings, New Orleans 1991, p265–270

13. W. Polhemus and C.B. Langford, "The Integration of Helmet Mounted Sight and Display Technology with Automated Target Handoff Technology", Proceedings of National Aerospace and Electronics Conference, Ohio, May 1977, p. 962–967

14. Warren Robinett, 1992, "Synthetic Experience: A Proposed Taxonomy", paper presented at the International Summer Institute on the State of the Art in Computer Graphics and submitted to Presence, 16th February, 1992

15. Ivor E. Sutherland, 1968, "A head mounted three–dimensional display", 1968 Fall Joint Computer Conference, AFIPS Conference Proceedings, Vol. 33, p.757–764.

16. Jih-fang Wang , Ronald Azuma, Gary Bishop, Vernon Chi,John Eyles and Henry Fuchs, 1990, "Tracking a head mounted display in a room sized environment with head–mounted cameras", SPIE volume 1290 Helmet Mounted Displays II, 1990

17. Graham Warwick, Simon Elliott, Ari Egozi and Gilbert Sebon, 1993, "Looks Can Kill", Flight International, 3–9 February, 1993, p.33–36

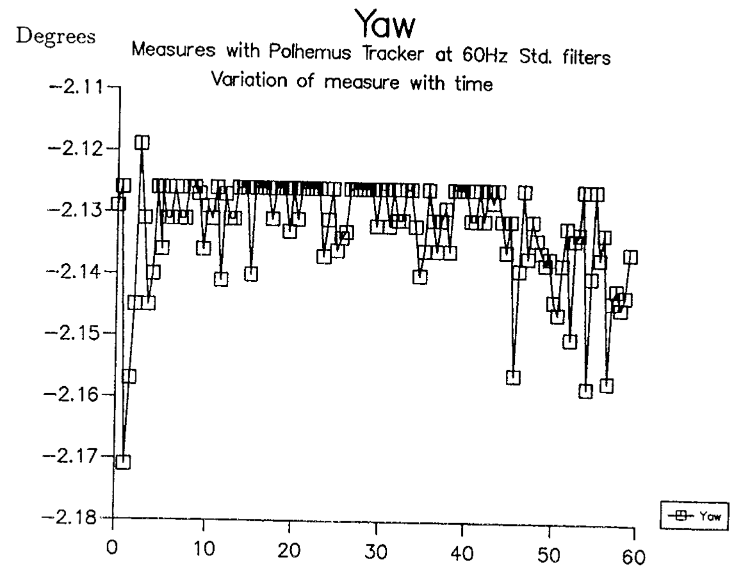


Figure 1.

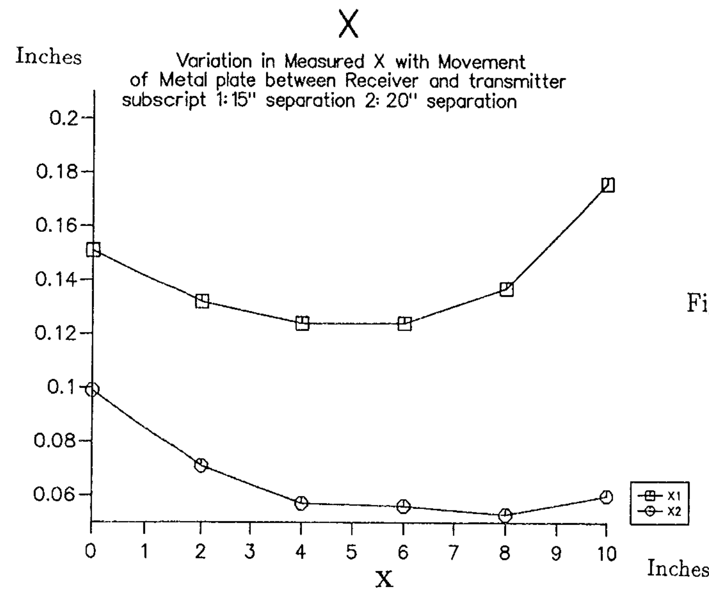


Figure 3.

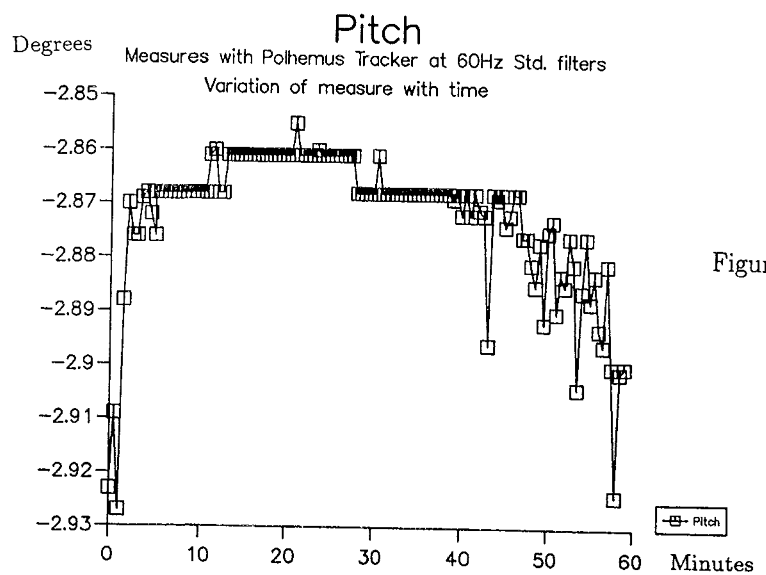


Figure 2.

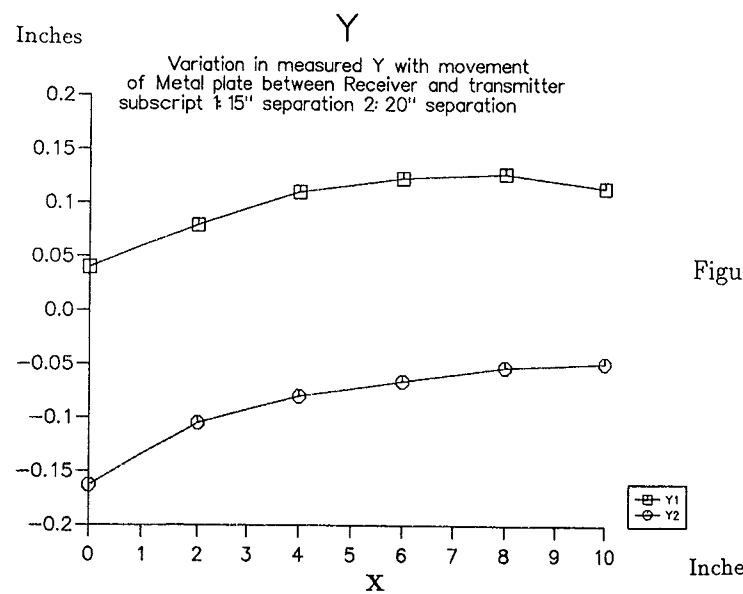


Figure 4.

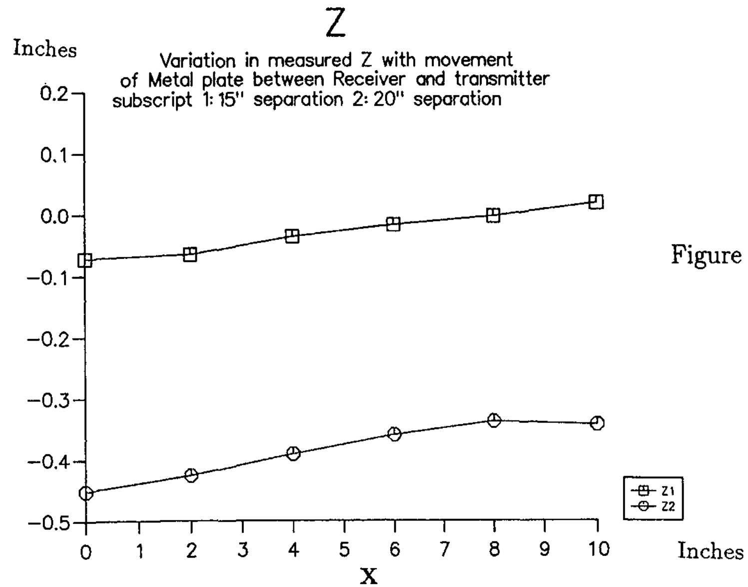


Figure 5.

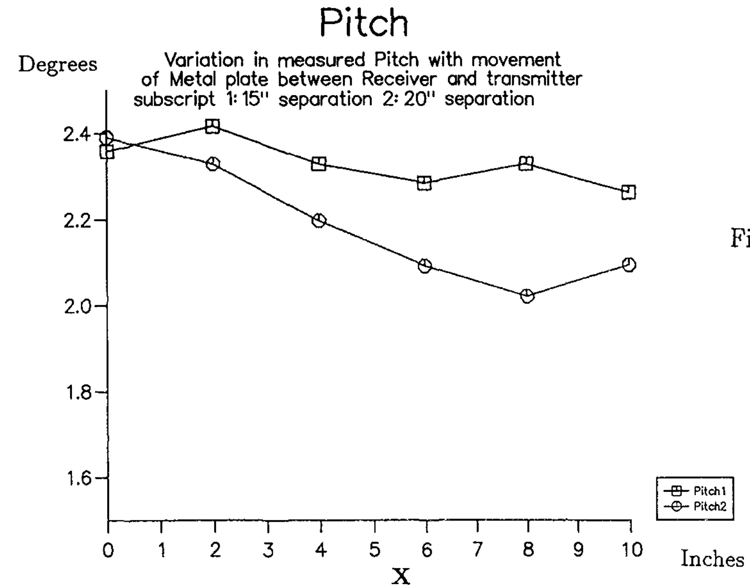


Figure 7.

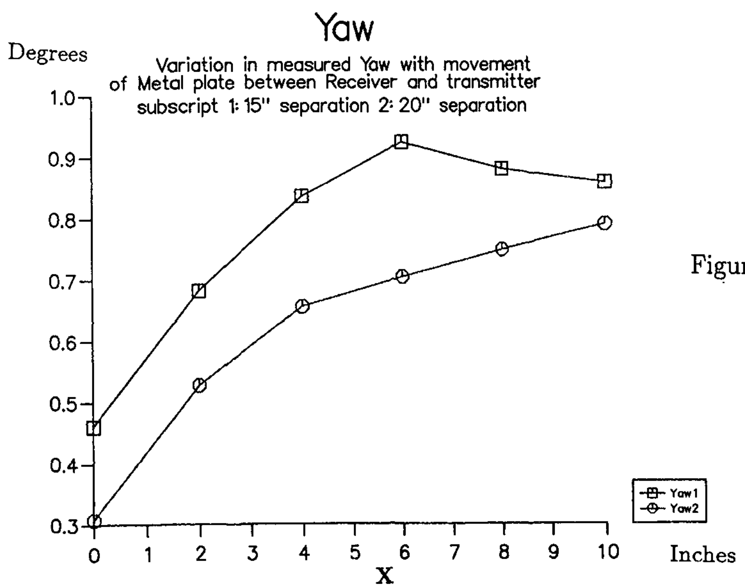


Figure 6.

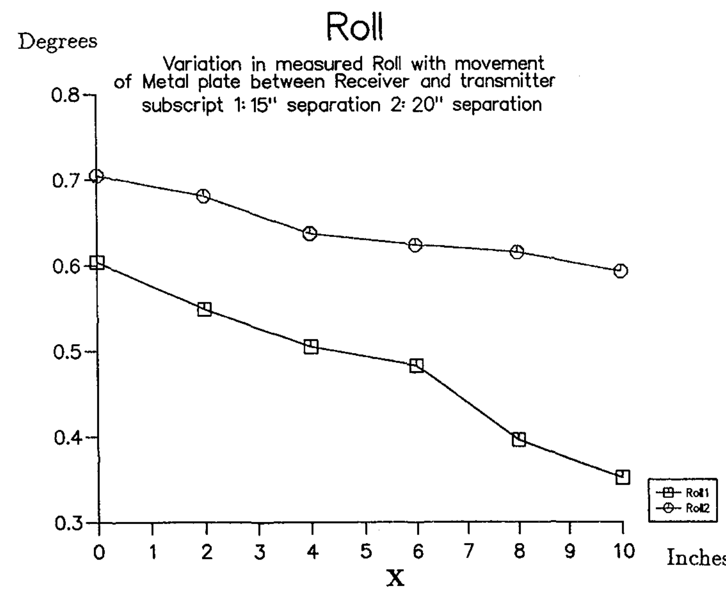


Figure 8.

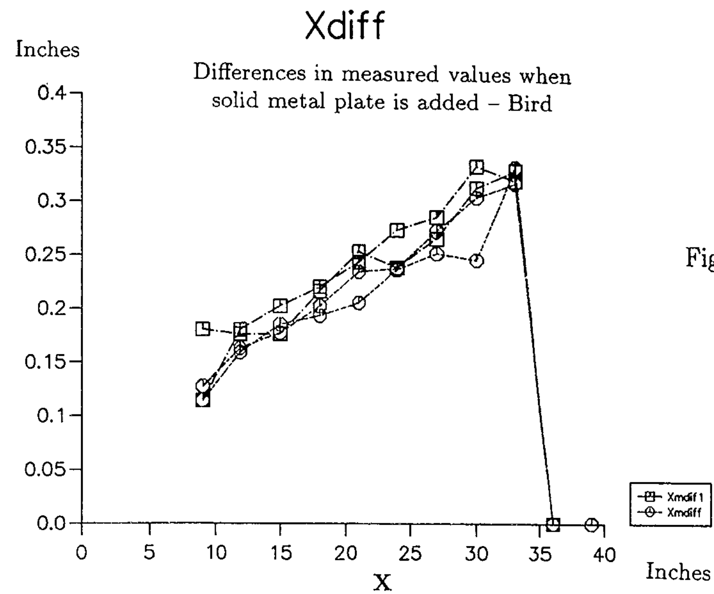


Figure 9.

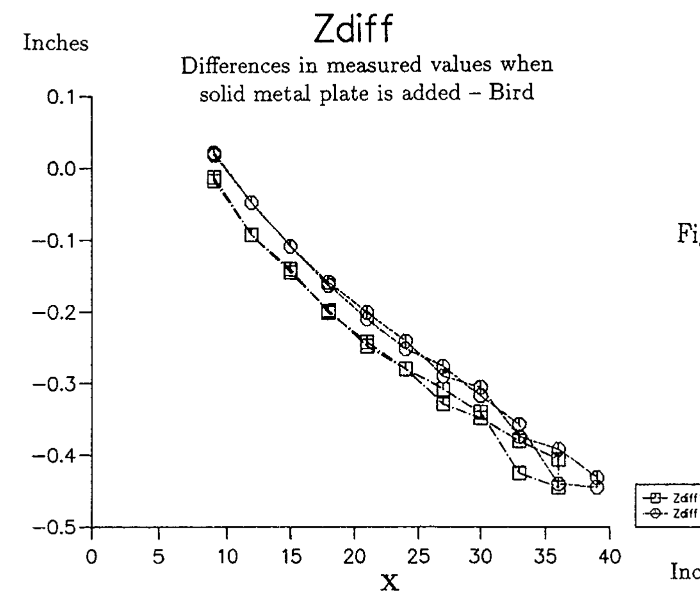


Figure 11.

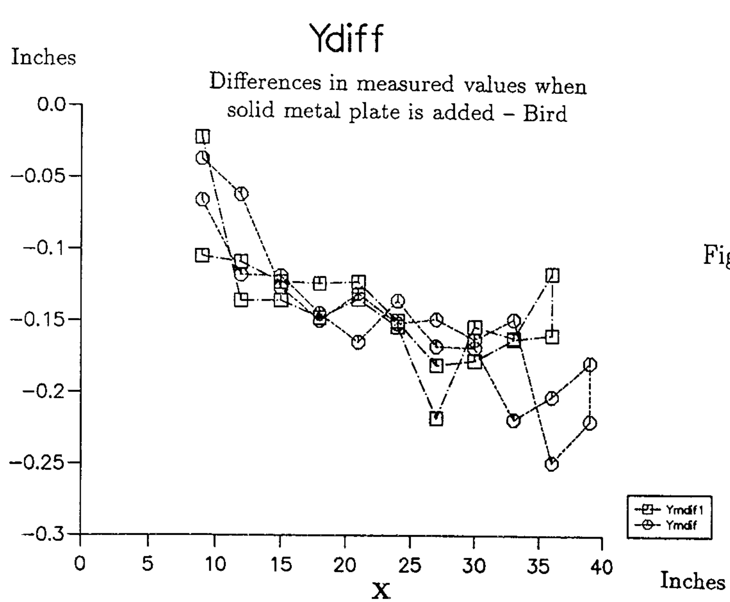


Figure 10.

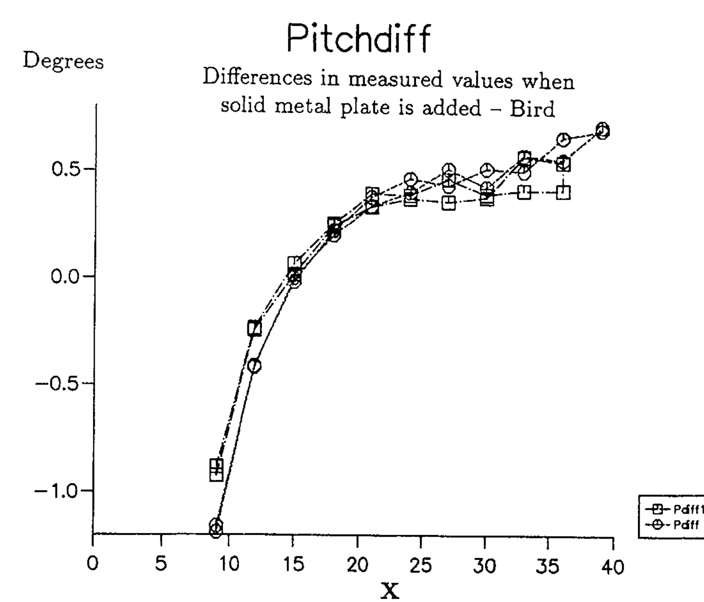


Figure 12.

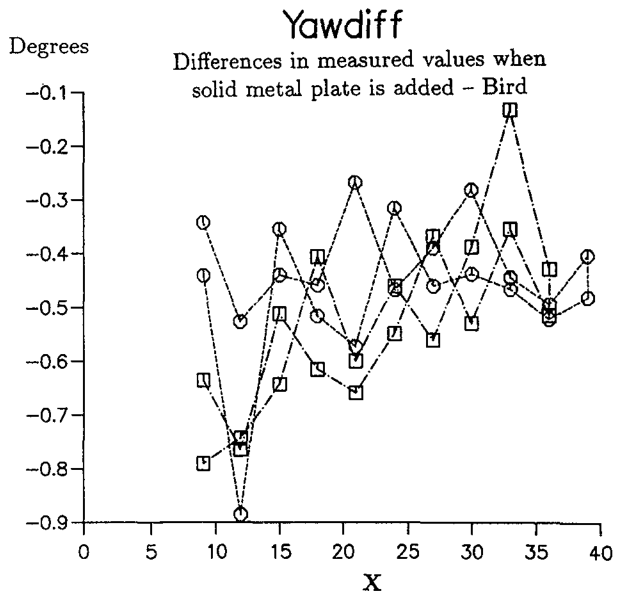


Figure 13.

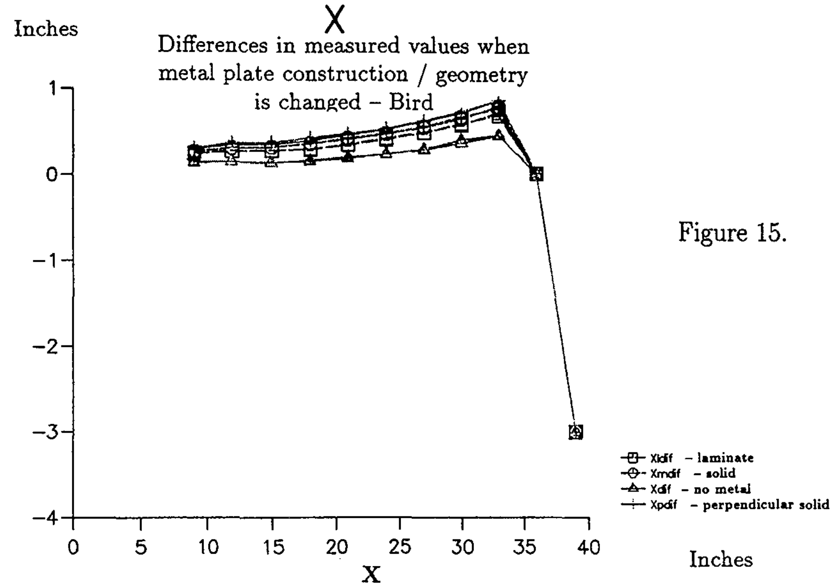


Figure 15.

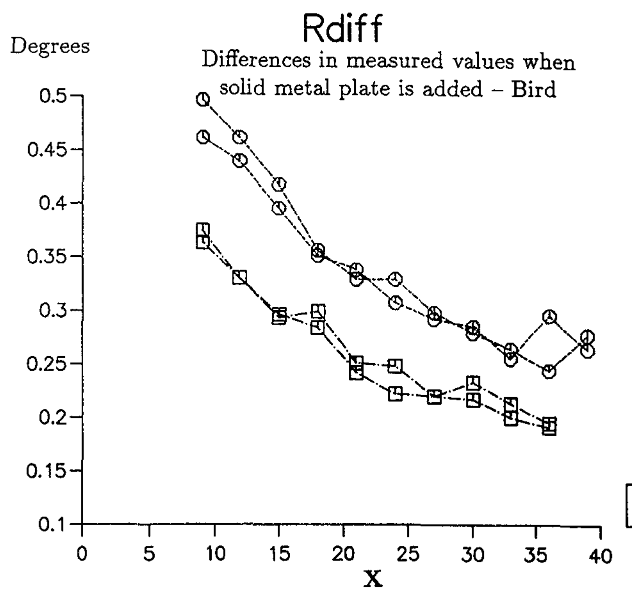


Figure 14.

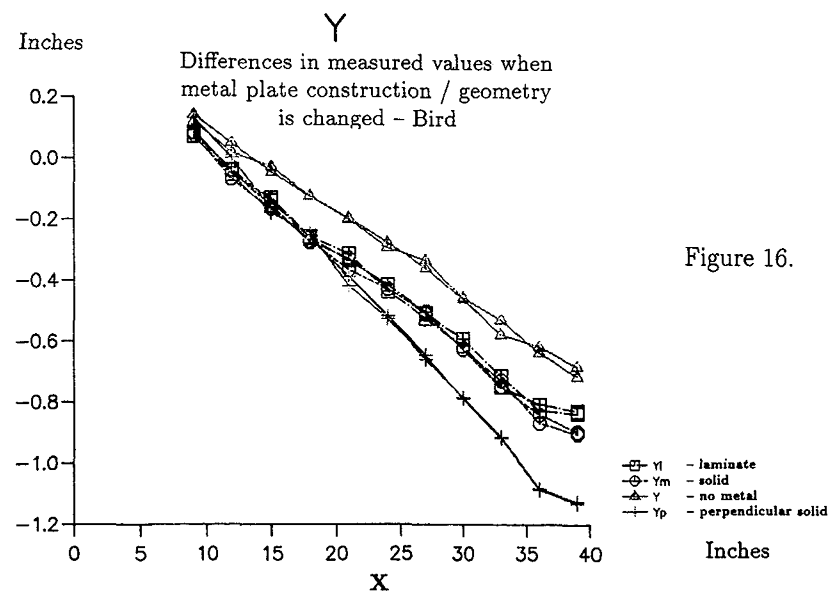


Figure 16.

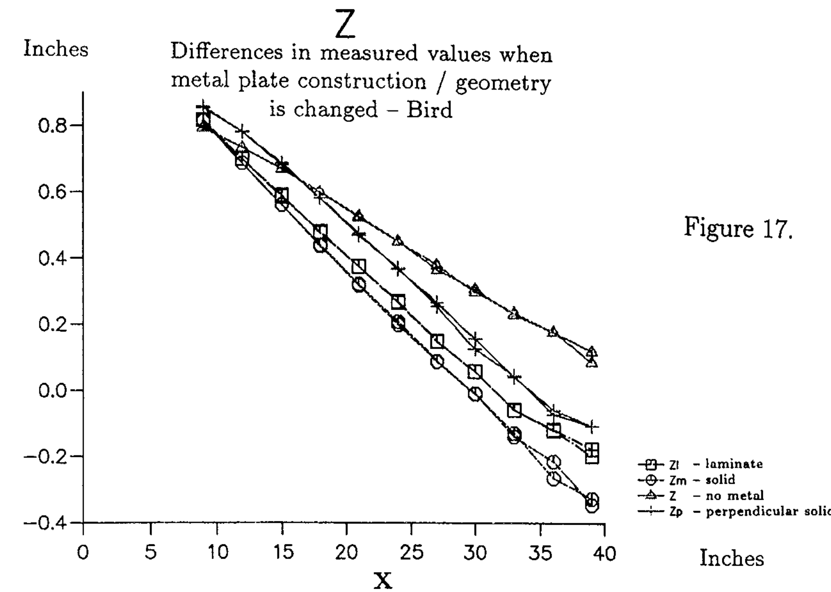


Figure 17.

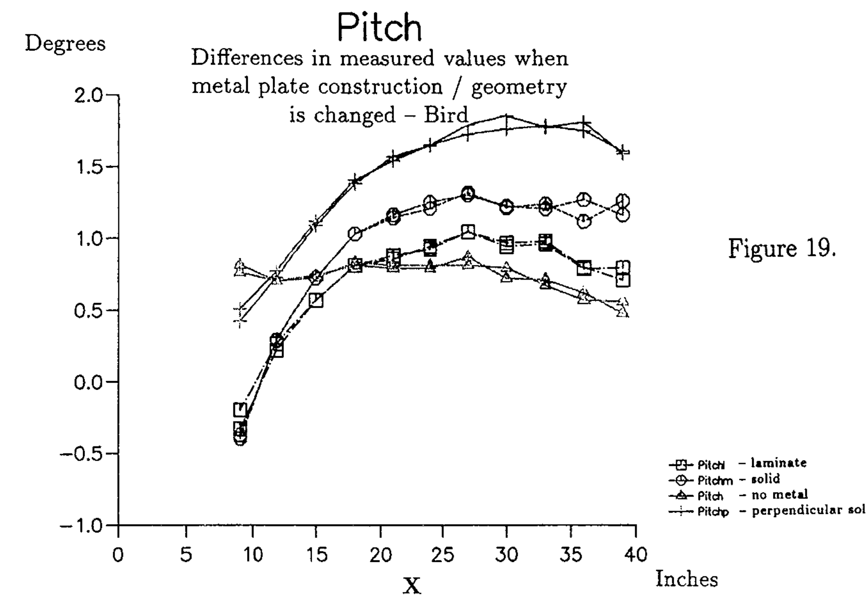


Figure 19.

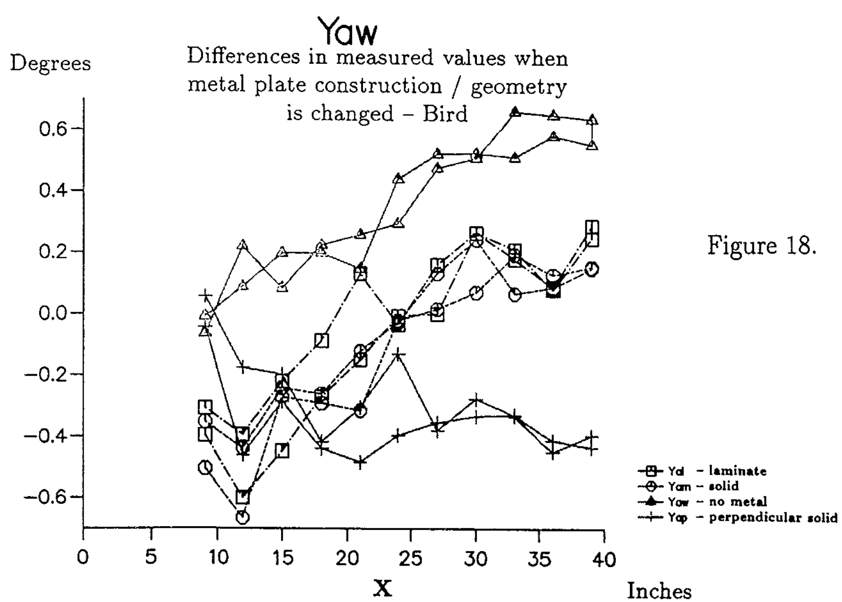


Figure 18.

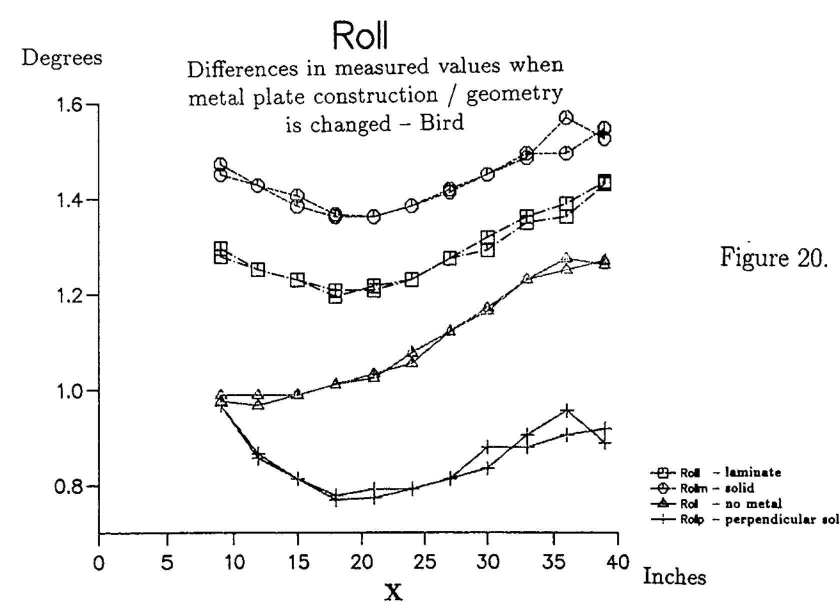


Figure 20.

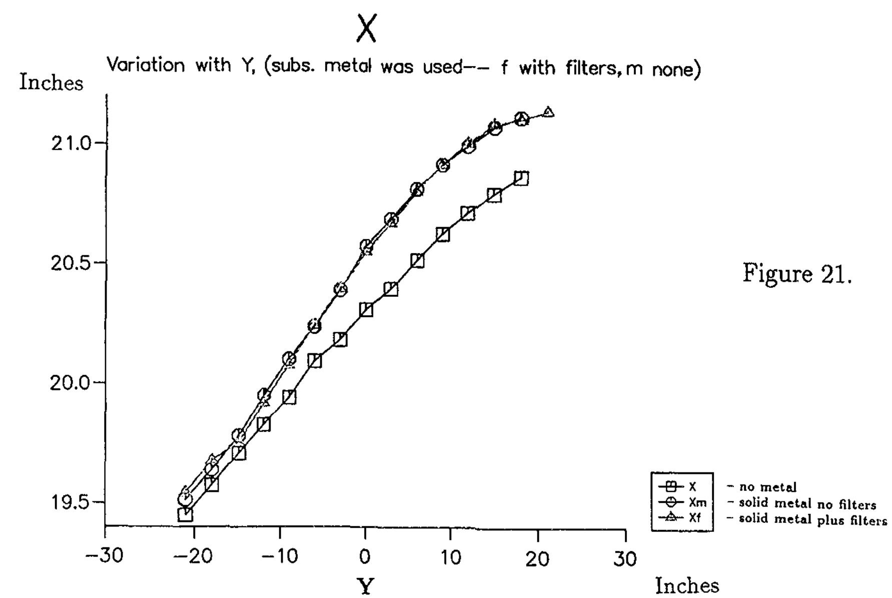


Figure 21.

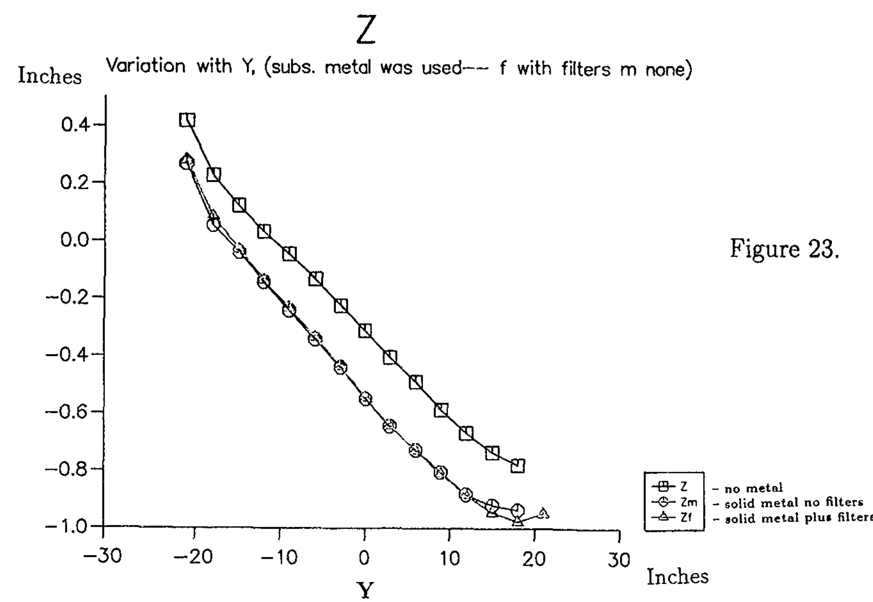


Figure 23.

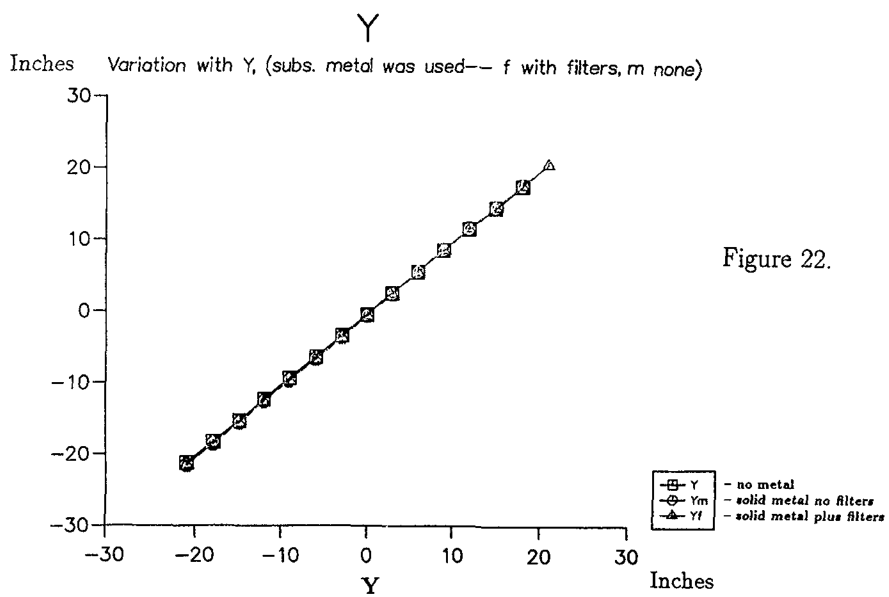


Figure 22.

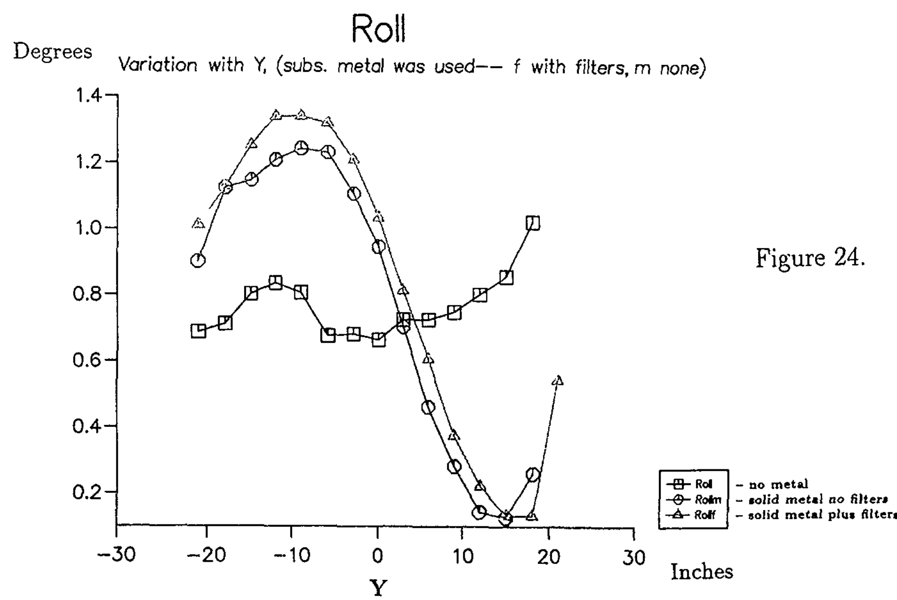


Figure 24.

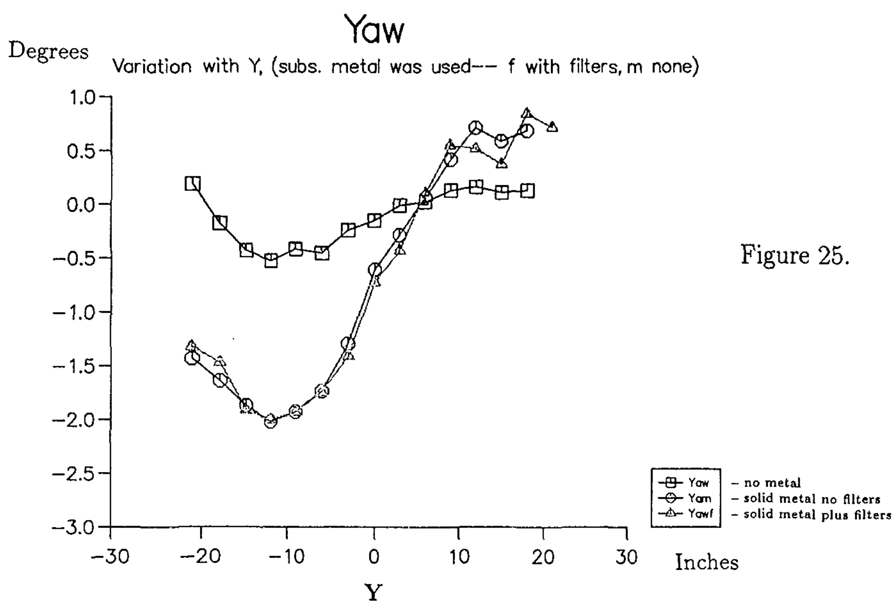


Figure 25.

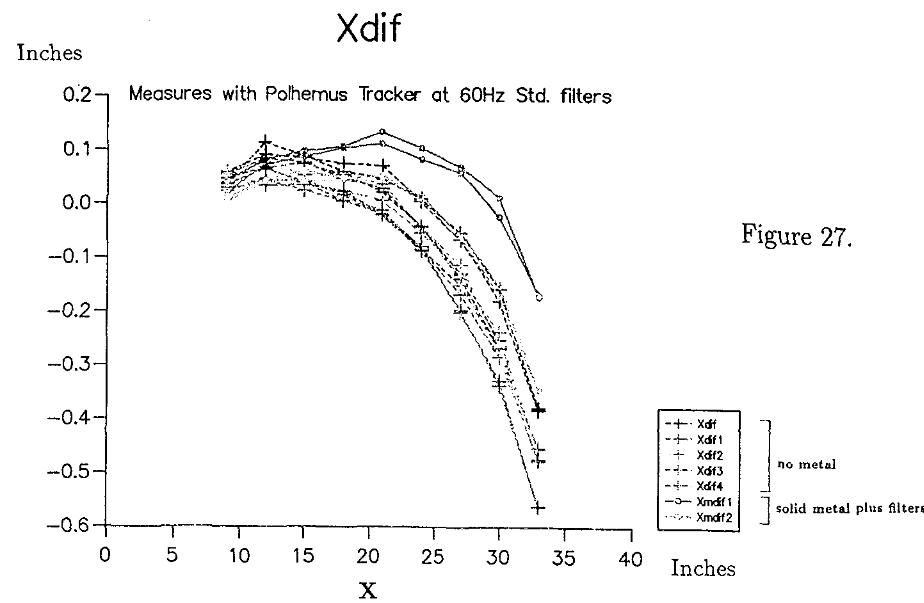


Figure 27.

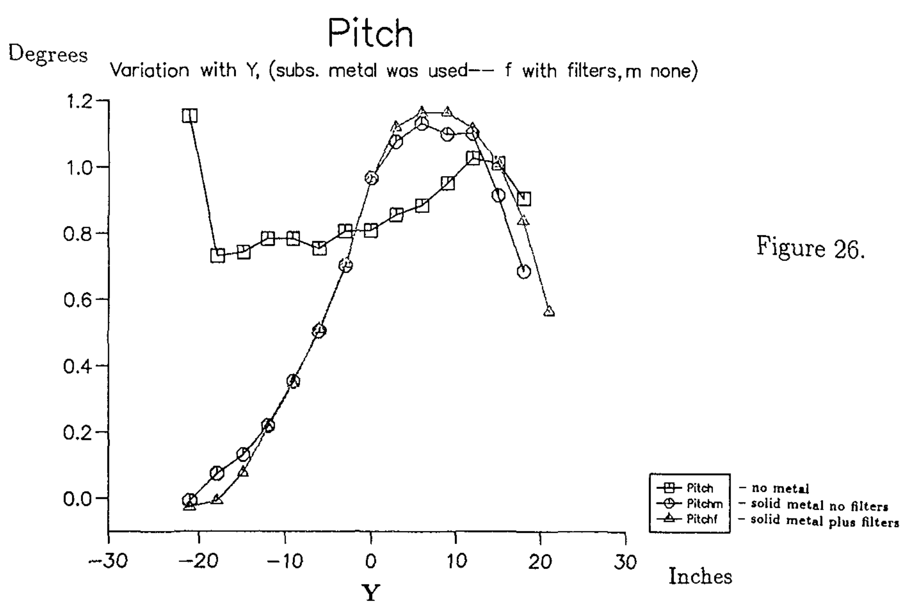


Figure 26.

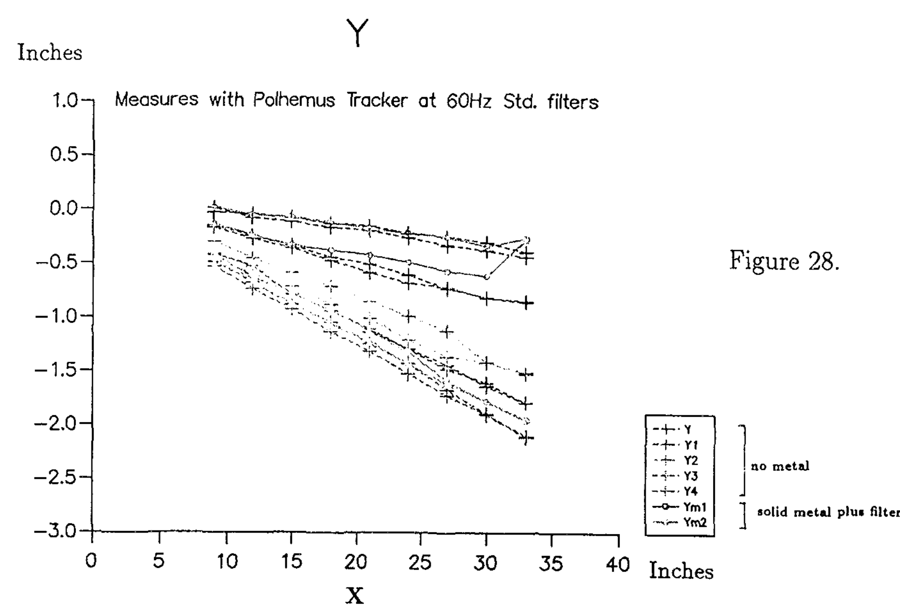


Figure 28.

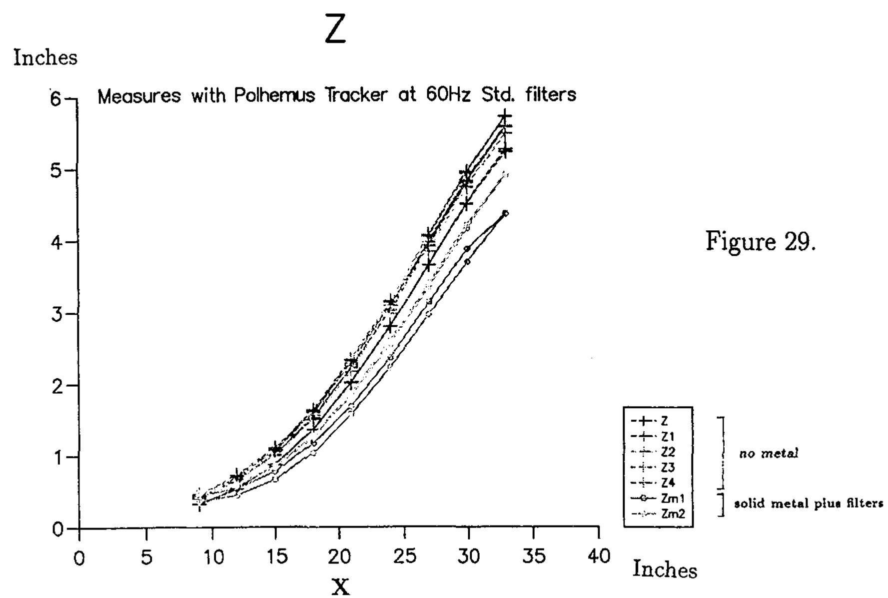


Figure 29.

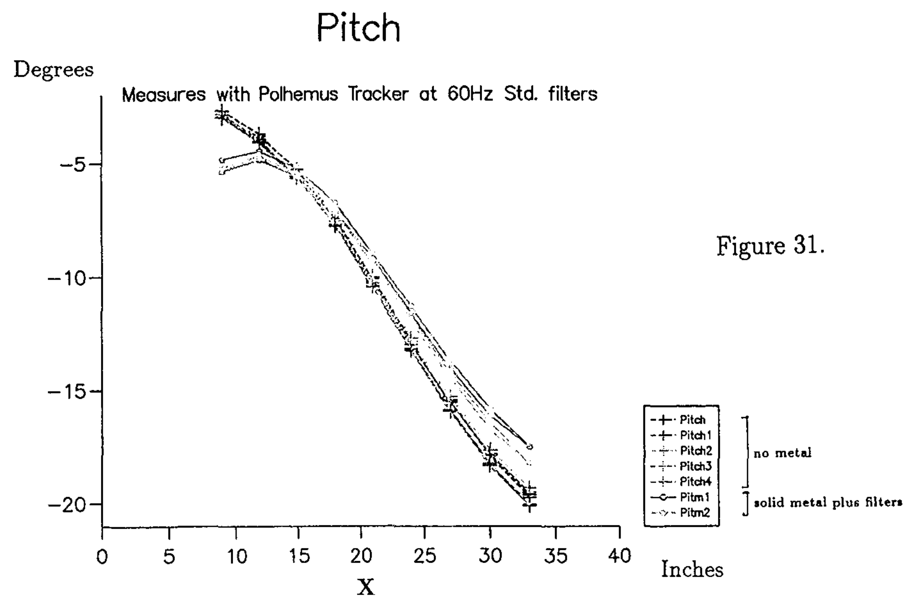


Figure 31.

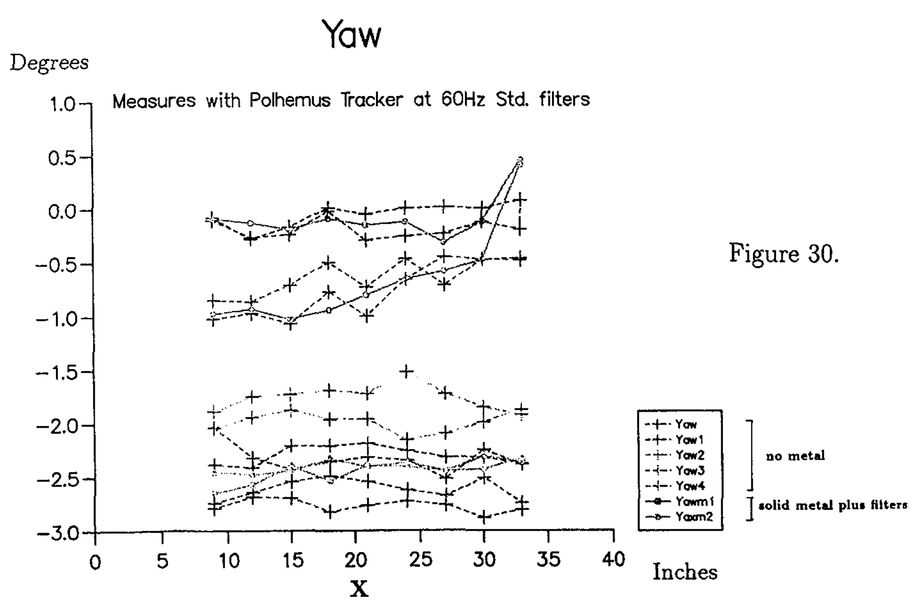


Figure 30.

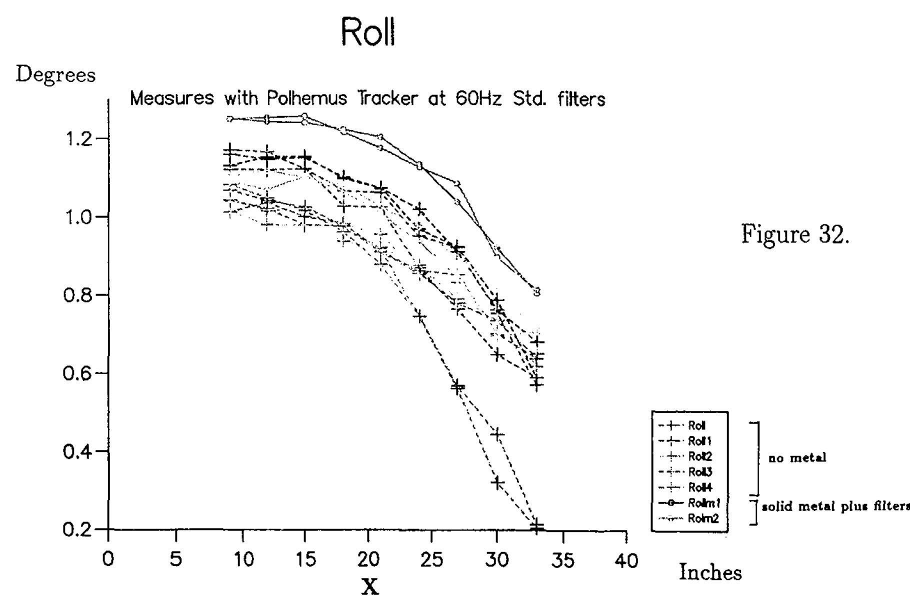


Figure 32.

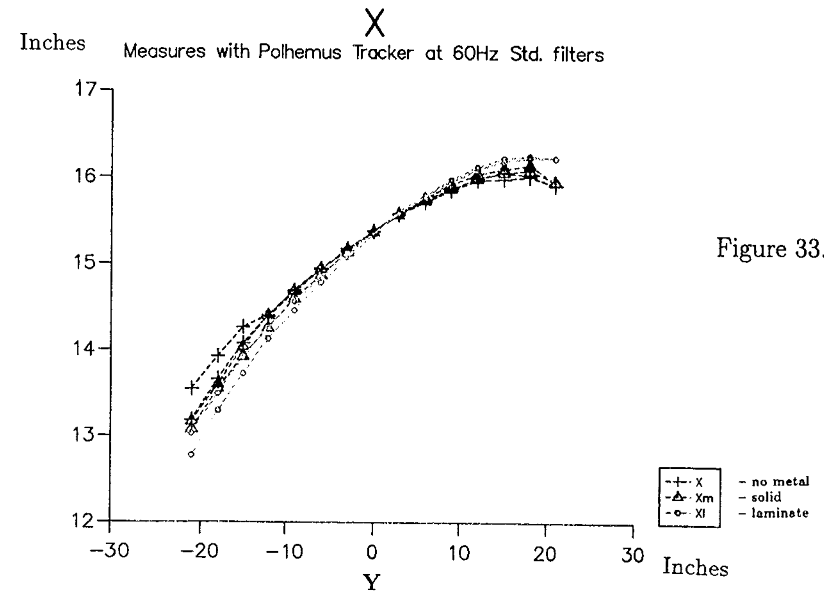


Figure 33.

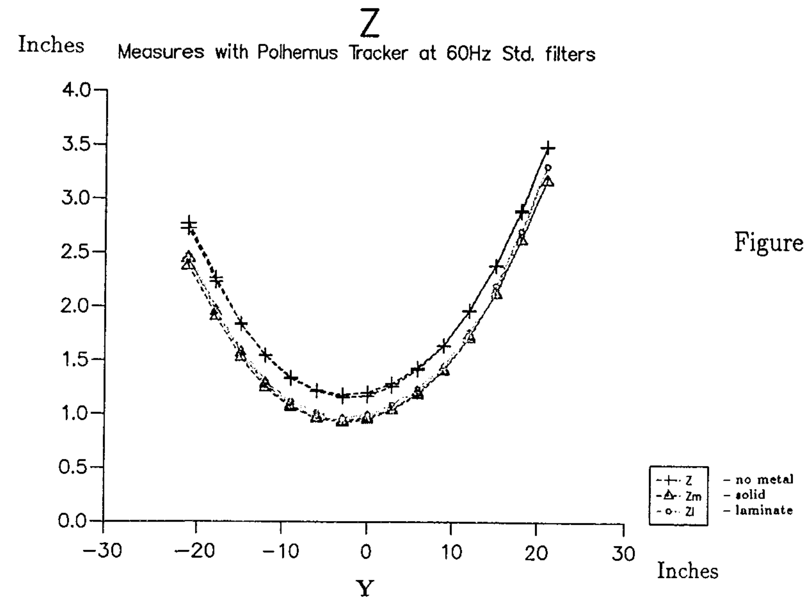


Figure 35.

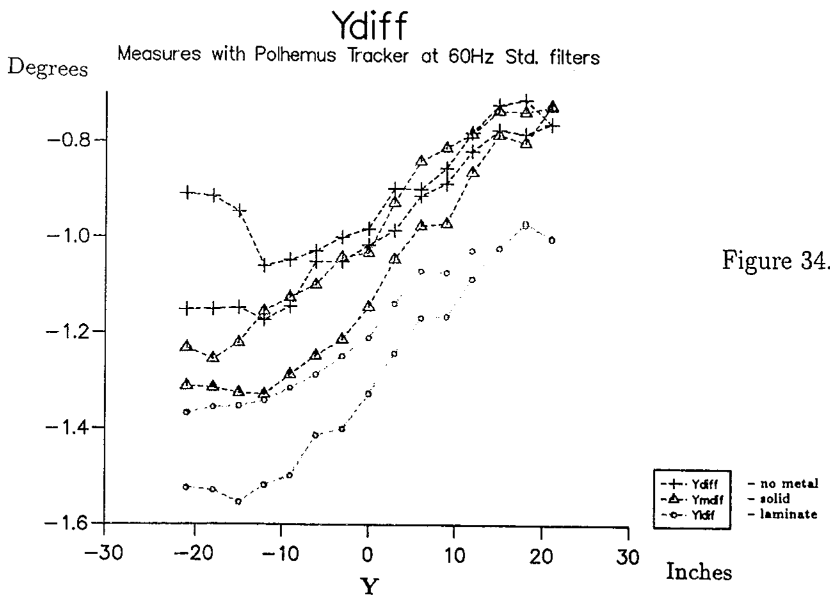


Figure 34.

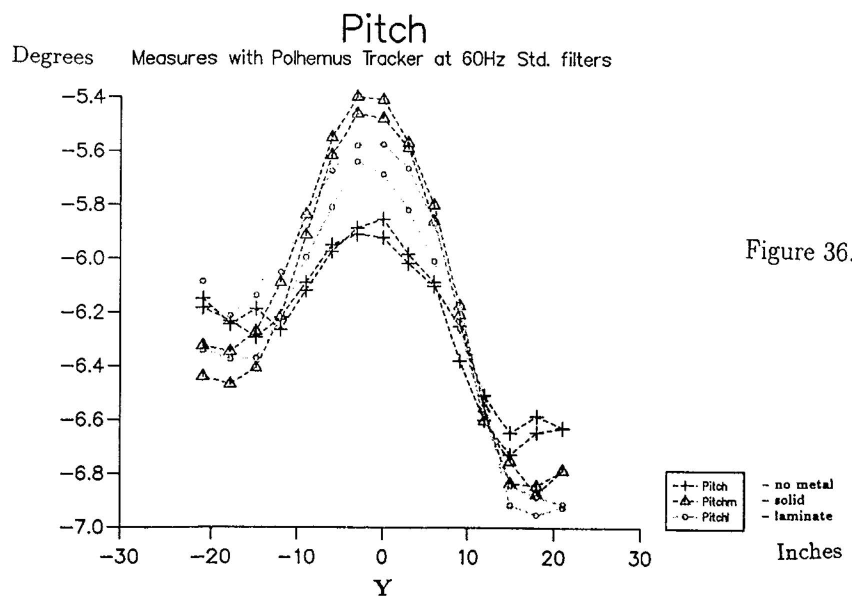


Figure 36.

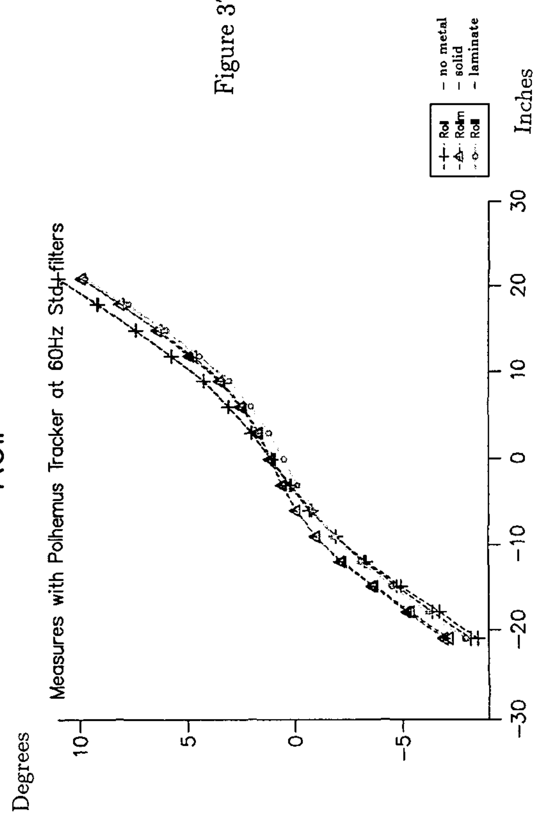


Figure 37.

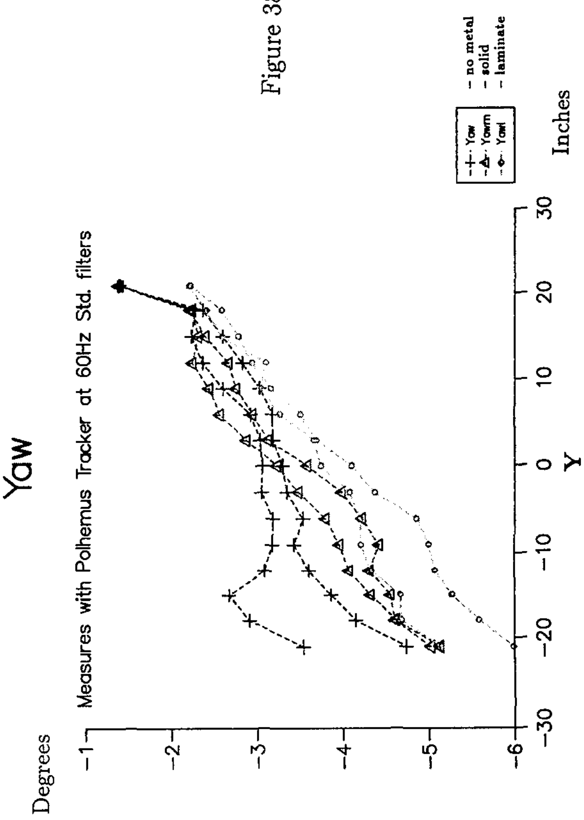


Figure 38.

ATTENUATING THE DISORIENTING EFFECTS OF HEAD MOVEMENT DURING WHOLE-BODY ROTATION USING A VISUAL REFERENCE: FURTHER TESTS OF A PREDICTIVE HYPOTHESIS

B. D. Lawson
F. E. Guedry
A. R. Rupert
A. M. Anderson

Naval Aerospace Medical Research Laboratory
Acceleration Division
51 Hovey Road
Pensacola, Florida 32508-1046
USA

SUMMARY

Research has shown that when subjects are seated upright and asked to perform an earthward head movement in the dark during whole-body rotation, they find the head movement disorienting if it is preceded by prolonged rotation at constant velocity, but not if it is made during the initial acceleratory phase of rotation. The disorienting effects of a head movement after prolonged constant velocity rotation can be attenuated by providing a visual reference to the Earth before the head movement. However, humans may not respond to vestibular or optokinetic stimulation the same way for different planes of motion. We tested the disorienting effects of an earthward head movement during rotation about a vertical axis to see if the attenuating effect of a visual reference would be altered. Some subjects were tested while lying on their side and some while lying on their back. Subjective reports concerning head movements in the dark were similar to previous research, suggesting that an acceleratory stimulus in the plane of rotation will attenuate disorientation, regardless of the plane of rotation tested. Likewise, the visual reference attenuated the disorientation that is usually associated with a head movement following prolonged constant velocity rotation. However, the visual reference did not appear to exert as strong an attenuating effect as it had for subjects seated upright. The implication of this finding for the design of centrifuge-based flight simulators is discussed.

INTRODUCTION

It is well known that if an individual is rotated and performs head movements in an axis that is not parallel to the axis of whole-body rotation, he will report feelings of spatial disorientation and eventually experience symptoms characteristic of motion sickness. This type of stimulation is known as Coriolis cross-coupling (**CCC**). Guedry and Benson (1978) demonstrated that the antecedent presence of an acceleratory stimulus to the horizontal semicircular canals ameliorates the disorienting effects of an earthward head movement in roll during whole-body rotation in the z-axis while seated upright. Guedry (1978) hypothesized that the aftereffects of large-field optokinetic stimulation in the horizontal plane will similarly modify activity in the vestibular nuclei as though the horizontal semicircular canals had been stimulated directly, thus attenuating the effects of CCC. Consistent with this explanation, his subjects no longer found earthward roll head movements after prolonged constant velocity rotation to be as disorienting if they were preceded by viewing an earth-fixed visual reference.

Thus, for the experimental situations that have been observed so far, certain kinds of preexisting vestibular activity (Guedry and Benson, 1978) or visual activity (Guedry, 1978) will decrease the amount of disorientation that an individual experiences during CCC stimulation. This finding should be interesting to the designers of centrifuge-based flight simulators, because any simulator profile that requires prolonged angular velocities about the central rotation

Presented at an AGARD Meeting on 'Virtual Interfaces: Research and Applications', October 1993.

axis coupled with angular motion of the simulator trainee's head (or the simulator cabin) about other nonparallel axes, will generate disorienting CCC stimulation. For example, to simulate the "g-forces" a pilot would experience during a steep bank and turn, a gimbaled centrifuge cabin could be used which would swing out from the central axis of the centrifuge during rotation. However, a much higher rate of rotation would inevitably occur in the centrifuge than during actual flight, enhancing the CCC effect. In fact, calculations show that even a 300-foot radius centrifuge would still generate appreciable angular velocity in the pitch plane of the semicircular canals during some of the more vigorous simulations, such as accelerating to 9 G_Z within 9 sec. To use a centrifuge to properly simulate the forces present in high-performance flight operations without producing disorientation during head movements, it will be necessary to understand the combinations of visual and vestibular information that can and cannot be expected to attenuate the disorienting effects of CCC stimulation. Certain types of visual displays presented with a virtual interface may reduce disorientation and simulator sickness while others may worsen it.

The main goal of this study was to test the observations of Guedry and Benson (1978) and Guedry (1978) for head movements made during previously untested axes of bodily rotation. There are good practical and scientific reasons to do these further tests. From the practical standpoint, we expect that trainees in centrifuge-based flight simulator operations would be required to make head movements in all three axes to perform their simulator duties, and likewise, that simulator cabins would be rotated in various axes to simulate various profiles of aircraft motion. Therefore, it is operationally important when considering the feasibility of centrifuge-based simulator training to have an appreciation of the perceptual responses of the trainee undergoing simultaneous rotation in multiple axes.

From the scientific standpoint, testing other axes of rotation is fundamental to understanding the process of human visual-vestibular integration in three dimensions. It appears that a wide variety of perceptual and gaze responses to real (or perceived) whole-body rotation are different in different axes of rotation, and we might expect these differences to affect the extent to which the disorienting effects of CCC stimulation can be attenuated by antecedent visual or vestibular inputs (see Guedry, 1974; Young, Oman, and Dichgans, 1975; Fetter, Hain, and Zee, 1986; Guedry, et al, 1990). Another way in which responses to rotation in the yaw plane may differ with other axes of rotation is the extent of directional symmetry they exhibit within a given axis of motion. Although response to CW and CCW yaw rotation about the earth-vertical, z axis is basically symmetrical (Matsuo and Cohen, 1984), there is evidence from animal research that vestibulo-ocular responses may be directionally asymmetrical during rotation about an earth-vertical axis while lying on the side (Money and Scott, 1962; Money and Friedberg, 1964; Money, McLeod, and Graybiel, 1965; Collins and Guedry, 1967; Darlot, Lopez-Borneo, and Tracey, 1981; Matsuo and Cohen, 1984). Whether there is an asymmetry in the vertical vestibulo-ocular reflex of humans is less clear; however, it cannot be ruled out (Hixson and Niven, 1969; Guedry and Benson, 1970 and 1971; Baloh, Richman, Lee, and Honrubia, 1983). Researchers have also reported directional asymmetries in the ability to visually suppress vertical nystagmus, differences in the degradation of visual acuity depending on the predominant beat direction of nystagmus, and directional differences in optokinetic nystagmus, perception of vection, and postural reactions while viewing an optokinetic stimulus moving in pitch (Money, et al, 1965; Hixson and Niven, 1969; Guedry and Benson, 1970, 1971; Benson and Guedry, 1971; Barnes, Benson, and Prior, 1978; Guedry, 1970; Matsuo and Cohen, 1984; Young, Oman, and Dichgans, 1975; Lestienne, Soechting, and Berthoz, 1977).

The present study consisted of four experiments to test the ability of antecedent vestibular and visual information to attenuate the disorienting effects of an earthward head movement following whole-body rotation in the previously untested pitch and roll plane of the head. We also investigated the possibility of directional differences in the attenuating effects of antecedent vestibular and visual information within a given axis of rotation.

METHODS

Subjects: Sixty-four research volunteers participated in the four experiments described

in this paper, with 16 subjects participating in each experiment. The majority of subjects in this study were naval officers awaiting assignment into flight training. Most of the subjects had little or no previous experience in rotating experiments, flight simulators, or acrobatic flight.

Materials: A Stille-Werner rotating chair and controller were used for testing. The original testing chair was replaced by a rotating litter with a hinged headrest, which could be triggered manually to allow for the head to drop passively 20 deg below the horizontal. The head movement was partially damped by padded stops and by the padded surface under the subject's head. An open framework made of PVC tubing was mounted on the litter and permitted the subject to view the interior of the earth-fixed experimental chamber through thin black vertical struts that were fixed relative to himself. (This was similar to the viewing conditions in Guedry, 1978.) The experimental chamber was a regular eight-sided polystyrene enclosure with a ceiling of the same material, whose walls each measured 4 ft wide and 8 ft tall. The white walls, ceiling, and visible portions of the floor were covered with black circular dots of 6 inches in diameter. The dots were placed pseudo-randomly such that an average of 45.5 dots were placed on each 4 ft x 8 ft panel; thus, 28.4% of the interior of the white chamber was covered with black dots. The distance from the subject's eye to the middle of any given wall was about 56.5 inches, and the distance from the subject's eye to the ceiling was about 60 inches. Thus, a given black dot would subtend not more than 6.1 deg of visual angle and not less than 5.7 deg, respectively. The rotating apparatus and the experimental chamber are shown in Figure 1.

Procedure: We positioned our subjects with their heads in the center of rotation and resting on the hinged headrest. In experiments 1A and 2A, subjects were positioned with their right sides down. In experiments 1B and 2B, they were tested while lying on their backs. The bodily positions assumed by subjects and the conditions for each experiment are shown schematically in Figures 2 and 3.

Subjective Measures: We asked the subjects in each experiment to compare the two types of head movements using paired-comparisons, and tell us which one they perceived to be more 'abnormal.' In experiments 1A and 1B, subjects were given four separate opportunities to compare a head movement made in the dark <u>immediately</u> after accelerating up to a constant dwell velocity **(ACC HM)** to a head movement performed <u>after 1 min</u> of rotation at constant velocity **(CONST HM).** In experiments 2A and 2B, they similarly compared a head movement performed after 1 min of rotation at constant velocity <u>in the dark</u> **(DK HM)** to a head movement made in the dark after 1 min of rotation <u>while viewing the illuminated interior of the experimental chamber</u> **(LT HM)**. A stationary rest period was allowed after each head movement, to ensure that all feelings of disorientation and all cardinal symptoms of motion sickness had abated for 2 min before the next rotation sequence (and subsequent head movement) was initiated. The comparison between the first and the second head movement of any pair was usually separated by no more than a 3-min interval for most subjects. Subjects made two comparisons between the two types of head movements (for a total of four head movements) on testing day 1, then took a 48-h rest before making two more comparisons on testing day 2. The four comparisons (of eight head movements) were made in random order and random rotation direction.

We chose the rotational velocity and the amplitude of the head movements to be mild enough to induce very little motion sickness (pilot study, n = 7). However, we carefully monitored any symptoms of motion sickness our subjects reported (Graybiel, Wood, Miller, and Cramer, 1968, Lawson, 1993). We also asked them to rate each of the head movements in terms of the magnitude of perceived 'disturbance' and 'disorientation' it evoked (Guedry and Oman, 1992; Guedry and Correia, 1971). Subjects were able to anchor their judgments of 'abnormality' by making several practice head movements <u>before</u> undergoing rotation, and a practice rotation that did <u>not</u> involve any head movement. These two baseline conditions were used as "normal" references for subsequent statistical analysis of ratings data. Immediately after making each head movement, subjects responded with the rating 'none', 'minimal', 'moderate', or 'major' to a variety of questions, which are briefly paraphrased below:

'DISTURBANCE':
1) "How immediately disturbing, abhorrent, or distressing was that head movement (versus baseline), apart from whether or not it was sickening?"
2) "How immediately startling, exciting, or arousing was that head movement (versus baseline), apart from whether or not it was sickening?"

'DISORIENTATION':
3) "How abnormal did that head movement feel (versus baseline) in perceived size and direction?"
4) "How abnormal did your body orientation seem versus baseline as a result of that head movement; i.e., did your body seem to tilt, tumble, dive, or otherwise move out of the horizontal plane of rotation?"

'MOTION SICKNESS':
5) "Please rate the magnitude of the following: nausea (including stomach awareness or discomfort), increased salivation, cold sweating, drowsiness, headache, dizziness, flushing/warmth, and skin pallor (rate by self observation in a mirror)."

In summary, subjects compared different pairs of head movements (CONST HM versus ACC HM and DK HM versus LT HM) to tell which felt more abnormal, and they also rated each head movement along a variety of dimensions during the baseline and rotating phases of the study. The median of four paired comparison judgments from each subject was included in an analysis of the Kendall coefficient of agreement, while median ratings of each head movement were analyzed with a Wilcoxon signed-ranks test.

RESULTS

Results are described separately for each of the four experiments, and are summarized in Table 1. Results for experiments 1A and 1B below describe paired comparisons between CONST HMs and ACC HMs, while results for experiments 2A and 2B refer to comparisons between DK HM and LT HM. We expected that CONST HM and DK HM will tend to evoke more abnormal perceptual effects.

Results of Experiment 1A (On Side in Dark): Fourteen of sixteen subjects were able to distinguish between ACC HM and CONST HM in all four of the paired comparisons they made, always calling CONST HM a more 'abnormal' experience. Conversely, no subject judged ACC HM as more abnormal than CONST HM in all four comparisons. Considering the pooled data from all 16 subjects, the Kendall coefficient of agreement $X^2 = 15.01$ for the median of all four comparisons, which was significant at $\alpha < .001$. In directional comparisons, 11/16 subjects did not tend to rate one direction of rotation as being more 'abnormal' than another ($X^2 = .14$ for median of two judgments, $\alpha > .70$ for $n = 16$).

Subjective ratings of each head movement during rotation were contrasted to the baseline conditions (HM without rotation and rotation without HM). Regardless of direction of rotation, subjects rated CONST HM as more disturbing, more startling, and more likely to induce abnormal perceptions of head and body orientation versus the baseline conditions. (Wilcoxon signed-rank tests: tie-corrected $z \geq 2.11$, $p \leq .038$). They were able to make these judgments even though they experienced little or no motion sickness as a result of CONST HM ($z = 1.62$ at $p = .10$). Conversely, no consistent perceptual effects were identified for ACC HM (i.e., no ratings significantly different from the baselines and independent of direction of rotation). In fact, ACC HM tended to produce somewhat less motion sickness than rotation without any head movement at all ($z \geq 2.12$ at $p \leq .034$). The best measures for distinguishing ACC HM from CONST HM were the subject's judgments of how disturbing and startling the head movement felt (see items 1 and 2 described in "Methods"). These measures distinguished CONST HM from the baselines regardless of rotation direction ($z \geq 2.11$ at $p \leq .034$), and showed no difference for ACC HM versus the baselines or for one baseline condition versus another.

Results of Experiment 1B (On Back in Dark): Of a total sixteen, six subjects judged CONST head movement as a more 'abnormal' experience than ACC HM in all four comparisons, while no subject found ACC HM more abnormal in all four comparisons. The Kendall coefficient of agreement $X^2 = 11.39$ for median of four comparisons was significant at $\alpha < .001$ for $n = 16$. 11/16 subjects did not tend to rate one direction of rotation as being more

'abnormal' than another (X^2 = .25 for median of two judgments, α > .50 for n = 16).

Regardless of direction of rotation, subjects usually rated the CONST HM as more disturbing, more startling, and more likely to induce abnormal perceptions of head and body orientation versus the baseline conditions. ($z \geq 1.89$, $p \leq .059$ for all ratings compared). They also tended to experience slight motion sickness as a result of CONST HM ($z \geq 2.18$, $p \leq .029$ for every contrast except CW rotation CONST HM versus CW rotation without head movement, where z = 1.81 at p = .07). Robust and consistent differences from the baseline conditions were not obvious during ACC HM. The best measure for distinguishing CONST HM from ACC HM was the subject's judgments of how disturbing the head movement felt (item 1 in "Methods"). This measure usually distinguished CONST HM from the baselines regardless of rotation direction (all $z \geq 1.89$ at $p \leq .059$), and showed no difference for ACC HM versus the baselines or for one baseline condition versus another.

Results of Experiment 2A (On Side with Visual Reference): Five of sixteen subjects judged the head movement made during constant velocity in the dark (DK HM) as a more 'abnormal' experience than the head movement made during constant velocity after viewing the visual reference (LT HM) in all four comparisons. No subject found the LT HM more abnormal in all four comparisons. The Kendall coefficient of agreement was X^2 = 3.13 for the median of all four comparisons, which did not quite reach significance (α = .08 for n = 16 subjects). However, 13/16 subjects found the DK HM to be more disorienting than the LT HM on the first opportunity they had to make a comparison, while 3 subjects could not distinguish any difference (X^2 = 5.29 at α < .02, n = 16). Eight subjects did not tend to rate one direction of rotation as being more 'abnormal' than another. Of the 8 subjects who did tend to rate one direction as more abnormal, 3 chose the CCW (pitch forward) direction on both opportunities to compare, and none chose the CW (pitch backward) direction on both opportunities. Overall, there was no significant directional preference (X^2 = 1.3 for median of two judgments, α > .20 for n = 16).

Subjective ratings of each head movement were highly variable, and consistent perceptual effects did not emerge from this analysis.

Results of Experiment 2B (On Back with Visual Reference): Of sixteen, six subjects judged the DK HM as a more 'abnormal' experience than the LT HM in all four comparisons, while no subject found the LT HM more abnormal in all four comparisons. The Kendall coefficient of agreement was X^2 = 6.25 for the median of all four comparisons, which was significant at α < .02 for all 16 subjects. Ten subjects found the DK HM more disorienting than the LT HM on their first opportunity to make a comparison (X^2 = 1.75 n.s. at α > .10, n = 16). Twelve subjects did not tend to rate one direction of rotation as being more 'abnormal' than another (X^2 = .05 for median of two judgments, α > .50 for n = 16). Subjective ratings of each head movement were variable, and no consistent perceptual effects emerged from this analysis.

DISCUSSION

The Potential Benefits and Drawbacks of a Centrifuge-based Flight Simulator: Flight simulators are a safe and relatively low-cost way of supplementing flight training. It is reasonable to expect that they will become a more integral part of flight training as they become increasingly more sophisticated and realistic. For example, if a virtual reality interface could be successfully coupled with a centrifuge-based motion platform, trainers would obtain more flexibility in the choice and the combination of acceleratory and visual information, making it possible to simulate a greater variety of high performance flight profiles with increased realism. However, it is important to recognize that disorientation and nausea will be evoked by certain combinations of centrifuge rotation with movements of the trainee's head (or his simulator cabin), and with particular visual stimuli. It is not sufficient to ignore this problem by simply allowing simulator trainees to adapt to these effects. Some situations that would cause disorientation in centrifuge-based simulators would probably not be disorienting during the actual flight operations being simulated (Gilson, Guedry, Hixson, and Niven, 1973), thus trainees might adapt in ways that are not

appropriate to real flight (Kennedy, Lilienthal, Berbaum, Baltzley, and McCauley, 1989). If centrifuge-based flight simulation is to be feasible without the risk of such negative transfer in flight training, it will be necessary to systematically identify simulator scenarios that produce disorientation only when it would occur during actual flight, and also to test the extent to which certain kinds of vestibular and visual information can attenuate these disorienting effects in cases where they are unavoidable during simulator training.

Head Movements in the Dark With and Without Acceleratory Information in the Plane of Body Rotation: The current study supports the idea that CCC stimulation will not be disorienting when it is immediately preceded by certain types of vestibular or visual information. Subjective reports concerning head movements in the dark were similar to previous research, suggesting that an antecedent acceleratory stimulus in the plane of rotation will attenuate feelings disorientation during CCC, regardless of whether the plane of body rotation tested is predominantly in the yaw, pitch, or roll plane of the semicircular canals. As summarized in Table 1, 100% of the subjects tested by Guedry and Benson (1978) judged the CONST HM as more disorienting than the ACC HM on their first (and only) opportunity to make a comparison, while 94% of the subjects in experiment 1A (on side in dark) and 63% of the subjects in experiment 1B (on back in dark) said the same thing on their first comparison. This antecedent acceleratory stimulus was particularly helpful in ameliorating feelings of disturbance, abhorrence, or distress that accompany CCC stimulation. It appears that the acceleratory stimulus is most helpful during CCC stimulation following rotation in the yaw axis of the head and least helpful following rotation in the roll axis of the head. The subjective reports following multiple exposures to CCC stimulation also indicate that the roll axis shows the weakest results of the three. However, the differences between the roll and pitch axes of rotation should not be over-emphasized, since the ratings that subjects made concerning each head movement indicated that they did not feel greatly abnormal perceptual effects during CCC stimulation in <u>either</u> the roll <u>or</u> the pitch axes of rotation. It is likely that the mild CCC stimulus chosen for this study is noticeably disorienting in the yaw plane, but only moderate effects exist in the roll and the pitch planes. This interpretation is consistent with the predominant feelings of body motion that subjects tend to perceive in each case. A subject sitting upright in the dark and rotating for a prolonged period at some constant velocity in the vertical axis tends to feel either a forward or a backward pitching sensation (depending upon the direction of rotation) if he executes a rightward head movement earthward. During normal circumstances, the detection of even the slightest <u>real</u> pitching motion while seated upright would be quite alarming and might be interpreted as a potential fall outside of the base of support. It would also require immediate postural compensation. On the other hand, a subject lying on his side or on his back making an earthward head movement during similar circumstances will tend to perceive rotation about his own longitudinal z-axis, as if he were simply rolling over in bed. The detection of a moderate motion of this type while lying securely restrained to a platform should not be quite as alarming or require the same postural adjustments.

Head Movements After Prolonged Rotation With and Without a Preceding Earth-fixed Visual Reference: Providing subjects with a visual reference prior to CCC stimulation tended to attenuate the disorientation associated with head movement following prolonged constant velocity rotation. However, the visual reference often did not appear to exert as strong an attenuating effect as the antecedent acceleratory information had for the experiments conducted in the dark. Moreover, the extent to which a visual reference will ameliorate the effects of CCC seems to depend upon the subject's plane of body rotation prior to the head movement. Guedry (1978) found that, depending upon the subject's gaze instructions, 81-100% of subjects receiving CCC stimulation after rotation in the yaw plane of the canals found the DK HM to be more disorienting than the LT HM on the first (and only) opportunity they had to make the comparison (see Table 1). When subjects were positioned on their sides in the present study and rotated in the pitch plane of the canals (experiment 2A), 81% reported that a DK HM was more disorienting than a LT HM on their first opportunity to make the comparison.

However, when subjects were positioned on their backs and rotated in the roll plane of the canals (experiment 2B), only 63% reported that a DK HM was more disorienting than a LT HM on their first opportunity to make the comparison. Moreover, the percentage of subjects in experiments 2A and 2B who felt that the visual reference was helpful on all four of the opportunities they had to compare DK HM and LT HM was much lower. It appears that an antecedent visual reference is most helpful in attenuating the initial effects of a single CCC stimulus after yaw rotation and least helpful after roll rotation. The paired comparisons suggest that the attenuating effects of a visual reference during a single CCC stimulus are not necessarily maintained during repeated stimulation. It is possible that this is partially attributable to shifts in judging criteria over time or to the mild nature of the CCC stimulus employed.

Lack of Trends in Subjective Ratings Data: The results discussed above focus mostly on the paired comparisons subjects made between DK HMs and LT HMs, that is, on their choice of which type of head movement was more perceptually abnormal. Subjective ratings were also made separately for each head movement in terms of the magnitude of the effect of CCC stimulation. These ratings were inconsistent and less likely to show significant differences than the paired comparison data. The specific subjective aspect of the CCC stimulation (e.g., disturbance, disorientation, motion sickness) that was attenuated by the visual reference tended to vary from subject to subject. The overall trend was for median ratings to be higher (i.e., more abnormal) for DK HM versus head movement without rotation, but this was also the trend for LT HM versus head movement without rotation.

We draw three inferences from the collective trends in the paired comparisons and ratings data: a) although the particular subjective aspect of the CCC effect that is attenuated by visual stimulation will vary, subjects are nevertheless sensitive to an overall attenuation of the perceptual effects when they make paired comparisons; b) it appears that although a visual reference attenuates the effects of CCC stimulation, it does not abolish them altogether; c) the magnitude of the subjective ratings (especially for motion sickness) suggests that the particular CCC stimulus used in this study was a fairly mild one for the body orientations tested (i.e., on side and on back) and may not be as amenable to measurement via subjective ratings as it is by a more sensitive paired comparison.

Validation of Comparison With Past Research Findings: It is possible that the results of the current study are not directly comparable to the experiment of Guedry, 1978, since there were many differences in protocol (e.g., rotation profile and dwell velocity, amplitude and nature of the head movement, and the exact nature of the visual stimulus). This possibility was tested in a recent follow-up study (see experiment 3, Table 1). Subjects (n = 12) were seated upright and restrained by a lap belt and a bite plate which allowed for a 20-deg earthward head movement strictly in the roll plane of the canals (along with some lateral head translation towards the right shoulder). They accelerated in the yaw plane of canals at 3 deg/s^2 to 90 deg/s for 2 min, then decelerated at 3 deg/s^2. They compared a head movement after 1 min of rotation in the dark (DK HM) to a head movement made (in the dark) after rotating for 1 min while viewing the illuminated interior of the polka-dotted chamber (LT HM). Most (92%) of these subjects found the DK HM to be more disorienting than LT HM on their first (and only) opportunity to make a comparison. These results further support the notion that a visual reference will be more helpful in ameliorating the effects of a single CCC stimulation for the original yaw rotations tested by Guedry.

Testing for Directional Asymmetries Within a Given Axis of Rotation: Subjects in the current study did not report any difference between the CW and CCW directions of rotation in any of the four experiments conducted. It is possible that such differences exist, but that they do not become apparent until higher rates of rotation are employed. It is also possible that directional differences were masked by the instructions to the subject to consider the different head movements collectively when comparing the different directions of rotation. For example, a subject might typically make one DK HM and one LT HM during CW rotation, then compared both of these head movements to both head movements (DK HM and a LT HM) during CCW rotation.

This pooled comparison is of particular concern when we consider that the primary focus of the subject's attention would have been on distinguishing differences between the DK HM and the LT HM, and only secondarily on distinguishing differences between head movement during CW versus CCW rotation.

These problems were addressed in four control experiments where the subject's only task was to compare head movements made in either direction of rotation. These experiments were identical to the four main experiments reported in this study, with some important exceptions. Subjects (total $n = 47$) were rotated (once in each direction) up to a constant velocity of 120 deg/s. Half of the subjects were rotated in the CW direction first, and half in the CCW direction first. In control experiments 1C and 1D, they compared a DK HM in the CW direction to a DK HM in the CCW direction. In control experiments 2C and 2D, they compared a LT HM in each direction of rotation. No prominent directional asymmetries were seen in the attenuating effects of antecedent vestibular or visual stimulation ($X^2 \leq 3.0$ at $\alpha \geq .09$, $n = 11\text{-}12$ in each control experiment).

The Role of Neck Kinesthesia: We should note that the control and appreciation of normal head movements is not solely achieved by the integration of visual and vestibular information alone, but is also dependent on the rich source of kinesthetic information available from the muscles and joints of the neck. Moreover, the time constant of integration that renders small differences in the duration of each subject's head movements negligible from the standpoint of the angular impulse delivered to the semicircular canals may indeed make a difference to the neck spindle receptors. To establish that effects reported so far result primarily from an outcome of the integration of visual and vestibular sources of information, we used a protocol in which the importance of active control of the neck musculature is diminished. Firstly, our subjects were rotated with their heads resting on a pad and strapped to a headrest. Secondly, they practiced making manually-triggered passive head movements until subject and experimenter both agreed that the head movement looked and felt like a passive drop. For example, when subjects were interviewed following (the first) experiment 1A, 15/16 said that all of the head movements they made felt passive. However, there is probably no such thing as a truly passive movement, especially of the head. In a further attempt to indirectly assess the possible role of neck information in this study, we ran an individual who presented as normal, except that he had suffered bilateral damage to the his labyrinths due to the administration of gentamycin 5 years earlier. This subject was unable to distinguish between DK HM and LT HM or between CONST HM and ACC HM during either on-side or on-back rotation. We conclude that while neck information is important to the control and appreciation of head movement, the present study has been adequately controlled such that neck kinesthesia is probably not sufficient *per se* to account for the effects we have described. Nevertheless, conditions that produce altered kinesthetic control of the neck can influence an individual's spatial awareness. Lackner and DiZio (1989, 1992) found that increasing the load to the head (with a mass) increased the disorientation and motion sickness evoked by CCC stimulation and by sinusoidal rotation under conditions where the vestibular stimulus was kept constant (Lackner and DiZio, 1989; 1992).

Summary: Our results collectively support the notion that the presence of an earth-fixed visual reference prior to a CCC stimulus tends to attenuate the disorientation an individual feels. We have extended the research of Guedry and Benson (1978) and Guedry (1978) to previously untested axes of rotation and found that the visual reference will tend to be most helpful in attenuating the effects of a single CCC stimulus when the predominant plane of stimulation of the vestibular system prior to the head movement is in yaw. Our observations indicate that while the designers of centrifuge-based flight simulator profiles may not be too troubled by differences in visual-vestibular integration in one direction of rotation versus another, they will have to address differences that exist for one axis of rotation versus another. It may also be necessary to take into account the amount of experience the trainee has had with the CCC stimulus. This makes the problem of centrifuge-based flight simulation a difficult one, indeed. Fortunately, group trends are consistent, that is, whenever the visual reference is less helpful in attenuating

the effects of CCC stimulation, it still appears to be of <u>some</u> benefit.

ACKNOWLEDGMENTS

This research was sponsored by the Naval Medical Research and Development Command under ONR grant number N0014/90/J/1549.

The views expressed in this article are this of the authors and do not reflect the official policy or position of the Department of the Navy, Department of Defense, not the U.S. Government.

Volunteer subjects were recruited, evaluated, and employed in accordance with the procedures specified in the Department of Defense Directive 3216.2 and Secretary of the Navy Instruction 3900.39 series. These instructions are based upon voluntary informed consent and meet or exceed the provision of prevailing national and international guidelines.

REFERENCES

Baloh, R. W., Richman, L., Lee, R. D., and Honrubia, V. 1983. The dynamics of vertical eye movements in normal human subjects. Aviation, Space, and Environmental Medicine, 54: 32-38.

Barnes, G. R., Benson, A. R., and Prior, J. 1978. Visual-vestibular interaction in the control of eye movement. Aviation, Space, and Environmental Medicine, 49: 557-564.

Benson, A. J., and Guedry, F. E. 1971. Comparison of tracking-task performance and nystagmus during sinusoidal oscillation in yaw and pitch. Aerospace Medicine, 42L 593-601.

Cohen, B., Matsuo, V., and Raphan, T. 1977. Quantitative analysis of the velocity characteristics of optokinetic nystagmus and optokinetic after-nystagmus. Journal of Physiology (London), 270: 321-344.

Collins, W. E., and Guedry, F. E. 1967. Duration of angular acceleration and ocular nystagmus from cat and man. I. Responses from the lateral and vertical canals to two stimulus durations. Acta Otolaryngolica, 64: 373-387.

Darlot, C., Lopez-Borneo, J., and Tracey, D. 1981. Asymmetry of vertical vestibular nystagmus in the cat. Experimental Brain Research, 41: 420-426.

Fetter, M., Hain, T. C., and Zee, D. S. 1986. Influence of eye and head position on the vestibulo-ocular reflex. Experimental Brain Research, 64: 208-216.

Gilson, R. D., Guedry, F. E., Hixson, W. C., and Niven, J. I. 1973. Observations on perceived changes in aircraft attitude attending head movements made in a 2-g bank and turn. Aerospace Medicine, 44: 90-92.

Graybiel, A., Wood, C. D., Miller, E. F., and Cramer, D. B. 1968. Diagnostic criteria for grading the severity of acute motion sickness. Aerospace Medicine, 39: 453-455.

Guedry, F. E. 1970. Vision during angular acceleration considered in relation to aerospace operations. In: Stahle, J. (Ed.), Vestibular Function on Earth and in Space. Oxford: Pergammon. 27-34.

Guedry, F. E. 1974. Psychophysics of vestibular sensation. In: Kornhuber, H. H. (Ed.), Handbook of Sensory Physiology. New York: Springer. Vol VI/2, 3-154.

Guedry, F. E. 1978. Visual counteraction of nauseogenic and disorienting effects of some whole-body motions: A proposed mechanism. Aviation, Space, and Environmental Medicine, 49 (1): 36-41.

Guedry, F. E., and Benson, A. J. 1970. Tracking performance during sinusoidal stimulation of the vertical and horizontal semicircular canals. In: Busby, D. E. (Ed.), Recent Advances in Aerospace Medicine. Dordrecht, Holland: Reidel. 276-288.

Guedry, F. E., and Benson, A. J. 1971. Nystagmus and visual performance during sinusoidal stimulation of the vertical semicircular canals. NAMRL-1131. Pensacola, FL: Naval Aerospace Medical Research Laboratory.

Guedry, F. E., and Benson, A. J. 1978. Coriolis cross-coupling effects: Disorienting and nauseogenic or not? NAMRL-1231. Pensacola, FL: Naval Aerospace Medical Research Laboratory. Aviation, Space, and Environmental Medicine, 49: 29-35.

Guedry, F. E., and Correia, M. J. 1971. Vestibular function in normal and in exceptional conditions. In: Masterson, R. (Ed.), Handbook of Behavioral Neurobiology. NY-London: Plenum Press, Vol. 1: 353-407.

Guedry, F. E., Norman, J. T., Turnipseed, G. T. and Rupert, A. H. 1990. Vestibular responses and motion sickness during pitch, roll, and yaw sinusoidal whole-body oscillation. NAMRL-1352. Pensacola, FL: Naval Aerospace Medical Research Laboratory.

Guedry, F. E., and Oman, C. E. 1992. Vestibular stimulation during a simple centrifuge run. NAMRL-1353. Pensacola, FL: Naval Aerospace Medical Research Laboratory.

Hixson, W. C., and Niven, J. I. 1969. Directional differences in visual acuity during vertical nystagmus. NAMI-1079. Pensacola, FL: Naval Aerospace Medical Institute.

Kennedy, M. A., Lilienthal, K., Berbaum, S., Baltzley R., McCauley E. 1989. Simulator sickness in U. S. Navy Flight Simulators. Aviation, Space, and Environmental Medicine, 60: 10-16.

Lackner, J. R. , and DiZio, P. 1989. Altered Sensory-Motor Control of the Head as an Etiological Factor in Space-Motion Sickness. Perceptual and Motor Skills, 68: 784-786.

Lackner, J. R., and DiZio, P. Altered Loading of the Head Exacerbates Motion Sickness and Spatial Disorientation. Presented at Aerospace Medical Association meeting, May 10, 1992, Miami, FL.

Lawson, B. D. 1993. Human physiological and subjective responses during motion sickness induced by unusual visual and vestibular stimulation. Doctoral Dissertation. NASA NGT 50306. Waltham, MA: Ashton Graybiel Spatial Orientation Laboratory, Brandeis University.

Lestienne, F., Soechting, J., and Berthoz, A. 1977. Postural readjustments induced by linear motion of visual scenes. Experimental Brain Research, 28: 363-384.

Matsuo, V., and Cohen, B. 1984. Vertical optokinetic nystagmus and vestibular nystagmus in the monkey: Up-down asymmetry and effects of gravity. Experimental Brain Research, 53: 197-216.

Money, K. E., and Friedberg, J. 1964. The role of the semicircular canals in causation of motion sickness in the dog. Canadian Journal of Physiological Pharmacology, 42: 793-801.

Money, K. E., and Scott, J. W. 1962. Functions of separate receptors of non-auditory labyrinth of the cat. American Journal of Physiology, 202: 1211-1220.

Money, K. E., Mc Leod, M. E., and Graybiel, A. 1965. Effect of vision on the two directions of vertical post-rotatory nystagmus. Proceedings of The Canadian Federal Biological Society, 8: 13.

Nutall, J. B. 1958. The problem of spatial orientation. Journal of the American Medical Association, 166: 431-438.

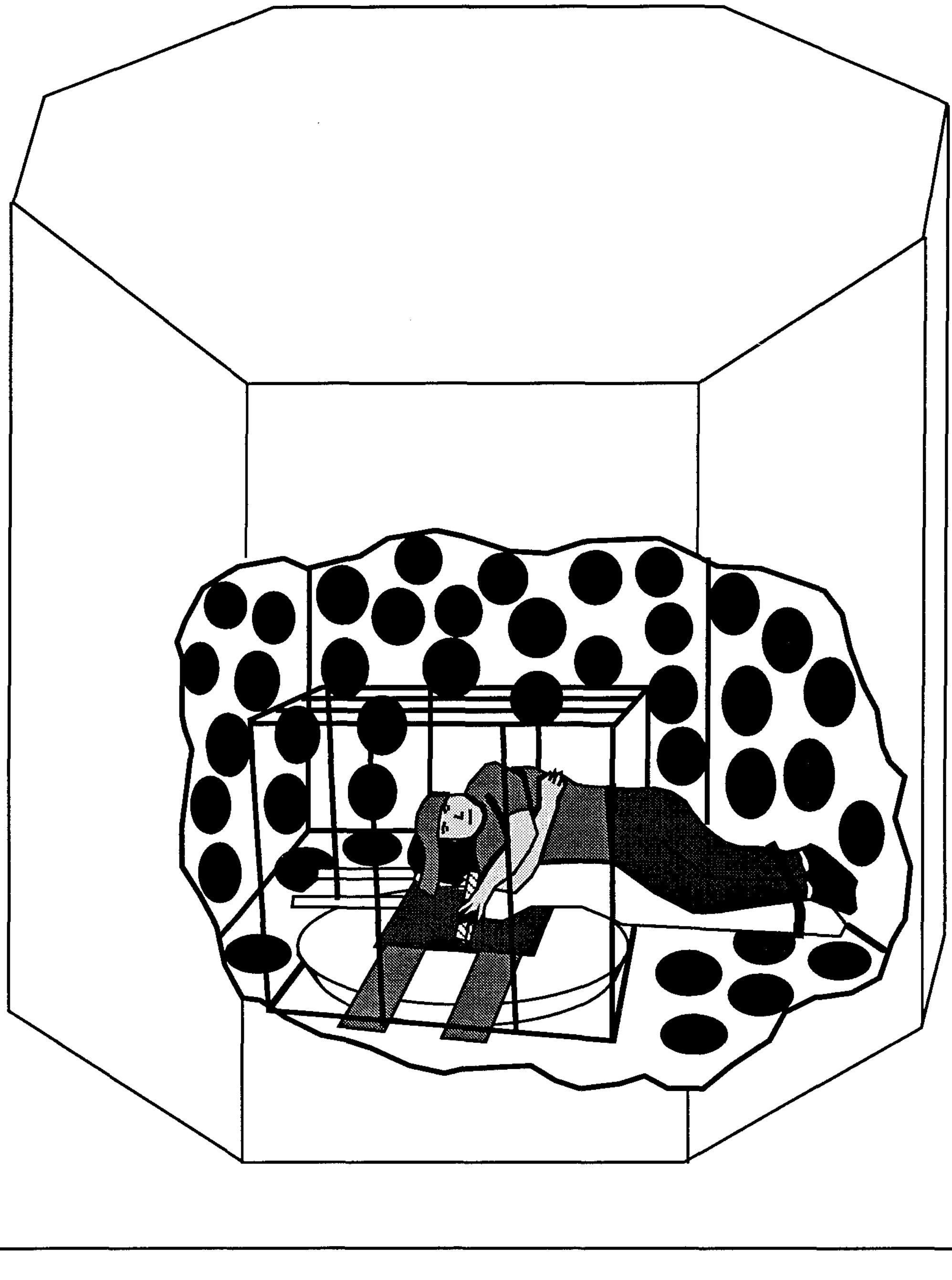
Raphan, T., Matsuo, V., and Cohen, B. 1979. Velocity storage in the vestibulo-ocular reflex arc (VOR). Experimental Brian Research, 35: 229-248.

Reason, J. T. and Diaz, E. 1970. The effects of visual reference on adaptation to Coriolis acceleration. FPRC Report 1303. London: Ministry of Defence (Air Force Dept.).

Young, L. R., Oman, C. M., and Dichgans, J. 1975. Influence of head orientation on visually induced pitch and roll sensation, Aviation, Space, and Medicine. 46: 264-268.

FIGURE 1

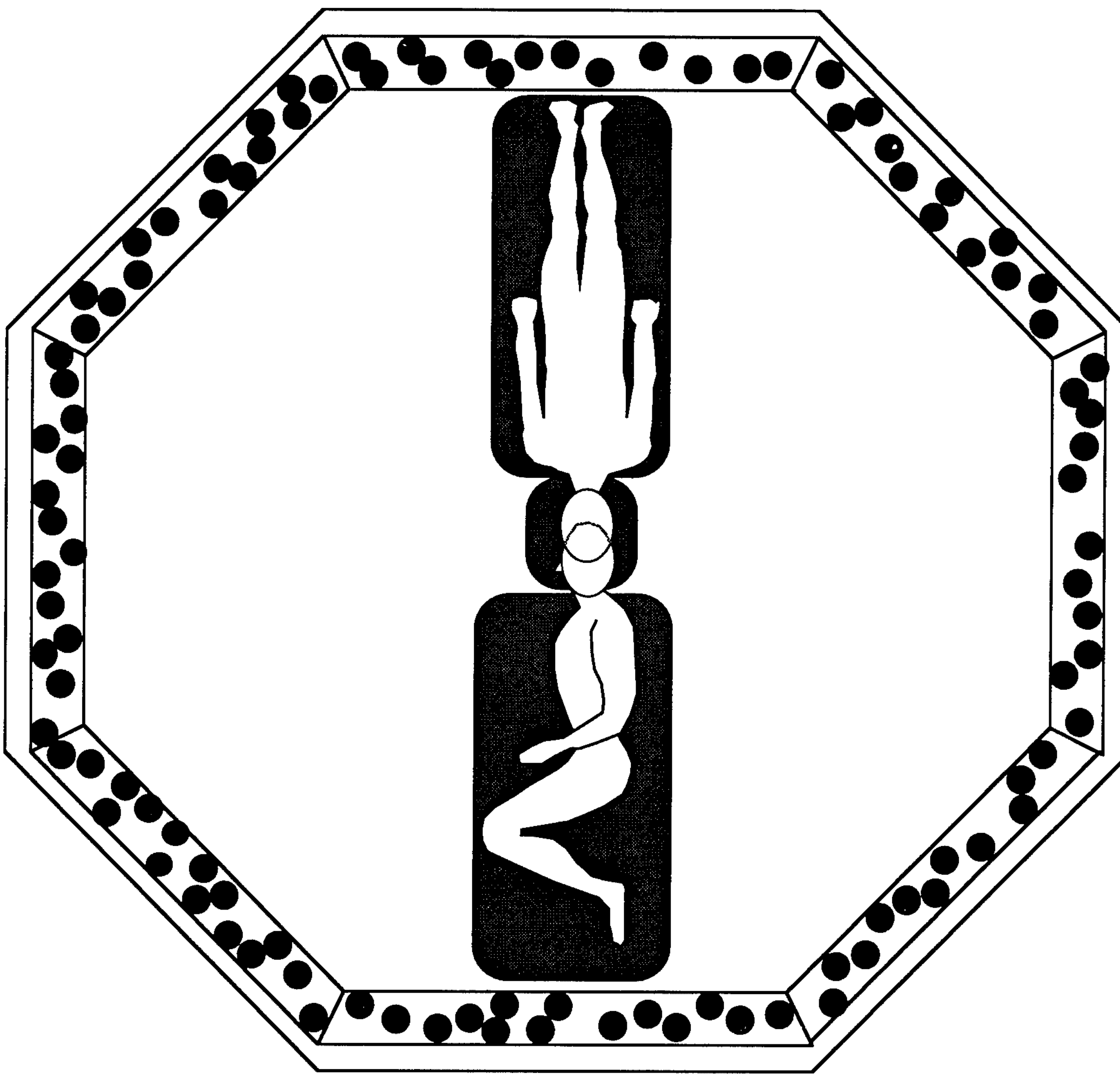
CUT-AWAY SIDE VIEW OF THE EXPERIMENTAL APPARATUS



The subject is shown lying on his right side in this case.
His head is resting on a hinged headrest in the center of rotation.
His left hand is on the triggering device which causes his head to drop passively.
Subjects were restrained at their heads, torsos, hips, and ankes during rotation.
In certain experiments, the subject gazed at the enclosure prior to making the head movement.

FIGURE 2

TOP-DOWN VIEW OF THE TWO BODY ORIENTATIONS TESTED



In experiments 1A and 2B subjects were rotated while lyin on their right sides.

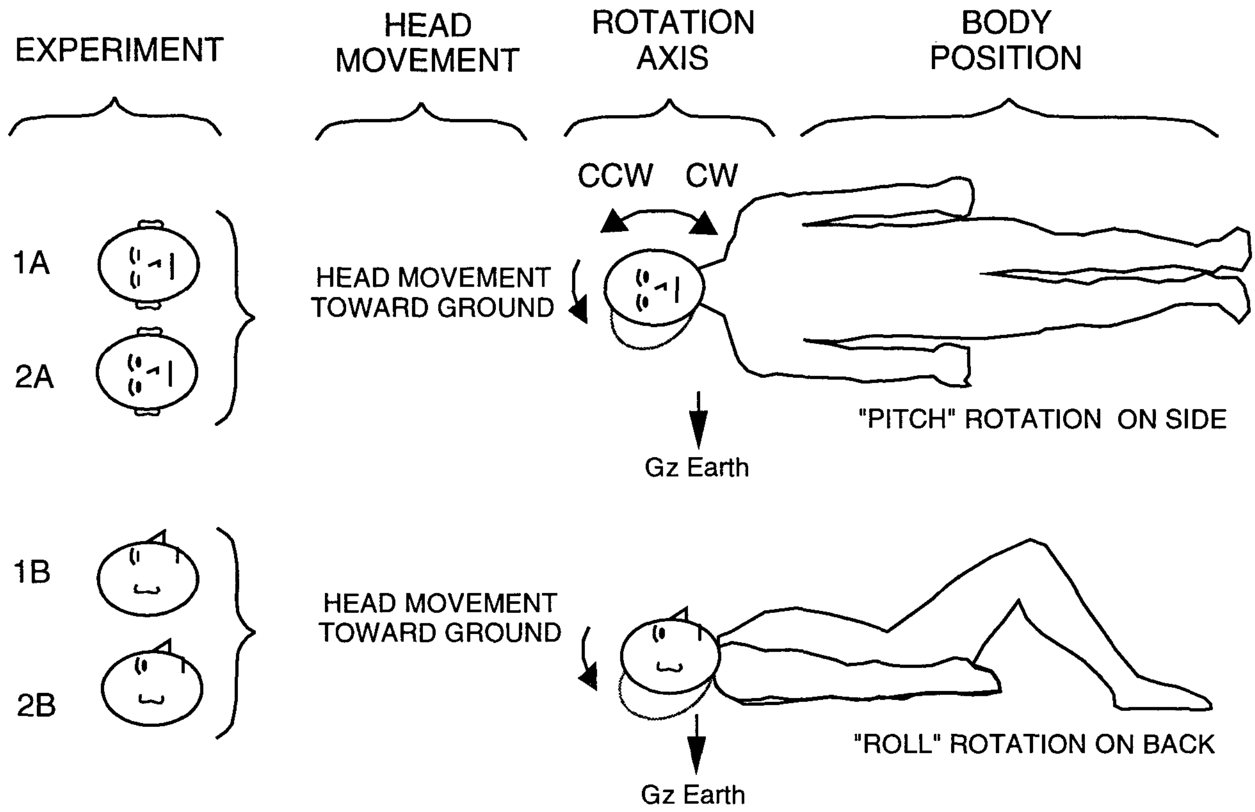
In experiments 1B and 2B they were rotated while lying on their backs.

Experiments 1A and 1B were run in darkness.

Experiments 2A and 2B were run with the chamber illuminated prior to each head movement.

FIGURE 3

DESCRIPTION OF THE 4 EXPERIMENTS THAT COMPRISED THE CURRENT STUDY



EXPERIMENT 1A:
Subjects (n=16) were positioned ont their sides and accelerated in darkness at 15 dg/s2 to a constant velocity of 90 dg/s for 2 mins, then decelerated at 3 dg/s2 to a stop. They compared an earthward head movement made <u>immediately</u> upon reaching constant velocity **(ACC HM)** to a head movement performed <u>after 1 minute</u> of rotation at constant velocity **(CONST HM)**. The **Coriolis cross-coupling (CCC)** stimulus was the same in either case, but the acceleratory information available to the vestibular apparatus was not (see Guedry and Benson, 1978).

EXPERIMENT 1B:
The protocol of experiment 1A was followed with subjects (n=16) lying on their backs.

EXPERIMENT 2A:
Subjects (n=16) were positioned on their sides and accelerated at 3 dg/s2 to 90 dg/s for 2 mins, then decelerated at 3 dg/s2. They compared a head movement after 1 minute of rotation <u>in the dark</u> **(DK HM)** to a head movement made (in the dark) after rotating for 1 minute <u>while viewing</u> the illuminated interior of the polka-dotted chamber (**LT HM**). The CCC stimulus was the same in either case, but visual information about the rotation was available to the vestibular nucleus during the LT HM (see Guedry , 1978).

EXPERIMENT 2B:
The protocol of experiment 2A was followed with subjects (n=16) lying on their backs.

TABLE 1

SUMMARY OF RESULTS FROM GUEDRY AND BENSON (1978), GUEDRY (1978), AND THE CURRENT STUDY

Experiment	Sample	Freq. During 1st Comparison	Freq. Unanimous During 4 Repeated Comparisons
PAST STUDIES:		Frequency reporting CONST HM more disorienting than ACC HM	
Upright in the Dark -- Guedry and Benson (1978)			
	$n = 12$	100 %	n.a.
		Frequency reporting DK HM more disorienting than LT HM	
Upright with Visual Reference -- Guedry (1978)			
"Schedule 1"	$n = 16$	81 %	n.a.
"Schedule 2"	$n = 6$	100 %	n.a.
PRESENT STUDY:		Frequency reporting CONST HM more disorienting than ACC HM	
Experiment 1A -- On side in dark	$n = 16$	94 %	88 %
Experiment 1B -- On back in dark	$n = 16$	63 %	38 %
		Frequency reporting DK HM more disorienting than LT HM	
Experiment 2A -- On side with a visual reference	$n = 16$	81 %	31 %
Experiment 2B -- On back with a visual reference	$n = 16$	63 %	38 %
Experiment 3 -- Upright with a visual reference	$n = 13$	92 %	n.a.

Some Side-effects of Immersion Virtual Reality

E C Regan
Special Psychology Research Group,
Army Personnel Research Establishment,
Ministry of Defence, Farnborough,
GU14 6TD, United Kingdom

SUMMARY

Virtual reality (VR) has become increasingly well-known over the last few years. However, little is known about the side-effects of prolonged immersion in VR. The main study described in this paper set out to investigate the frequency of occurrence and severity of side-effects of using an immersion VR system. Out of 150 subjects 61% reported symptoms of malaise at some point during a 20 minute immersion and 10 minute post-immersion period. These ranged from symptoms such as dizziness, stomach awareness, headaches, eyestrain and lightheadedness to severe nausea. Some research which has been conducted which attempted to identify those factors that play a causative role in the side-effects of the VR system is discussed. Finally, some areas for future research are highlighted.

1. BACKGROUND

Immersion VR has become increasingly well-known over the last few years. In an immersion VR system the user wears a headset which projects the virtual world, usually through two Liquid Crystal Displays (LCDs) which are mounted in the headset, and presents the illusion of actually being present in, or immersed in, the virtual world. Several companies, primarily in the UK and the US, now manufacture and supply VR systems, and the market for such systems is developing. With this developing market a whole host of applications for VR systems - ranging from military training to medical practice - have been suggested.

Many current simulators have side-effects on users. The most prevalent side-effect of modern simulators is 'simulator sickness' which occurs with many high performance simulators. Incidents of simulator sickness have been reported since 1957 when Havron & Butler (1) provided the first account of simulator sickness in aircraft simulators. Symptoms of simulator sickness are often similar to those of motion sickness, but generally affect a smaller proportion of the exposed population and are usually less severe. There is a possibility that similar sickness may occur with immersion VR systems. Furthermore, given other characteristics of immersion VR systems such as low resolution, display update lags, and full visual immersion in VR, side-effects such as visual problems and problems of disorientation and dizziness may be more likely to occur. However, at present there appears to be no documented literature concerning this.

The primary aim of the study which will be described next was therefore to document the frequency of occurrence and severity of side-effects of immersion in a VR system.

2. HARDWARE AND SOFTWARE

2.1 Hardware

A PROVISION 200 immersion VR system was used. This is a VR development platform based on Intel i860s and dedicated image generation processors. It can be expanded for multiple VR peripherals and multiple participants. A Virtual Research Flight Helmet was used to present the visual information. The Flight Helmet uses LCDs each with a resolution of 360 x 240 pixels. The field of view of the Flight Helmet is approximately 110° horizontally by 60° vertically. A 3D mouse was used for interaction with the system. This is a 6° of freedom pointing device that allows the user to move forward and backwards, pick up and manipulate objects. Both hand and head position were tracked using a Polhemus Fastrak tracking system.

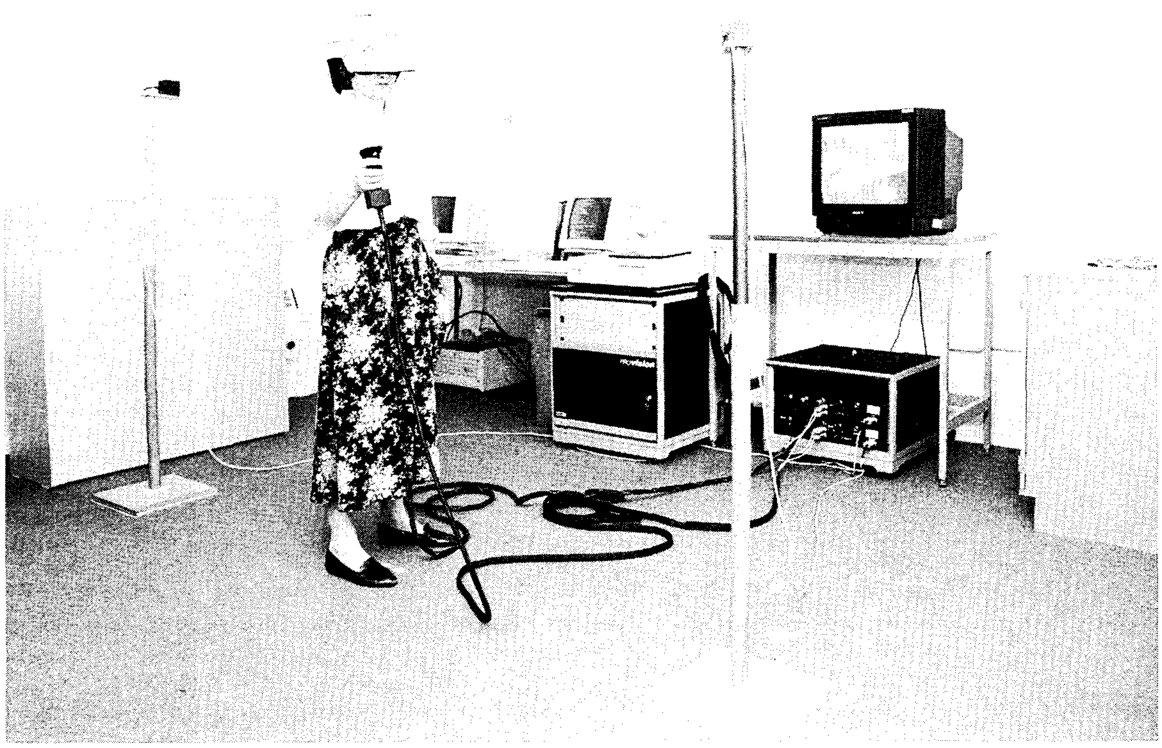
2.2 Software

For each subject the virtual world consisted of a corridor off which there were several doors leading into rooms. The subject was able to go into all of these rooms, and whilst in a room was able to interact with the objects in the room (for example, by picking them up and moving them). The rooms all contained different objects. One room, for example, contained a large chess board with pieces, and another contained a bar with bar stools and television with remote control unit. Each subject was given information about each room entered, and the ways in which the items in the room could be interacted with. Every attempt was made to ensure that all subjects underwent similar experiences in the virtual world.

The photograph below shows an individual immersed in virtual reality. The virtual world can be seen on the monitor.

Presented at an AGARD Meeting on 'Virtual Interfaces: Research and Applications', October 1993.

Figure 1 : The PROVISION 200 virtual reality system



3 . METHOD

One hundred and fifty subjects participated in the experiment. They consisted of 80 civilian subjects, 20 military subjects, and 50 firefighters. There were 106 male and 44 female subjects. Each subject was tested individually. The experiment lasted for approximately one hour per subject.

The subjects initially completed a 27 item symptom checklist (2), frequently used in simulator sickness research. This was also completed at the end of the immersion period.

Each subject was then immersed in the VR system for twenty minutes. Prior to the immersion the system was calibrated to the height of each subject, and the principles of the system and interaction with the system were explained to the subjects. A malaise scale was also described to the subjects and they were informed that they would be asked to rate themselves on this scale at 5 minute intervals during the twenty minute immersion period, and at 5 and 10 minutes post-immersion. The malaise scale had six categories as follows:-

1 = No symptoms
2 = Any symptoms, but no nausea
3 = Mild nausea
4 = Moderate nausea
5 = Severe nausea
6 = Being sick

A pre-immersion rating on the malaise rating scale was given by each subject and then the VR helmet was placed on the subject's head. The helmet was tightened and was then switched on. The stopwatch was started and the subject was told to proceed through the virtual world. After twenty minutes the helmet was switched off and removed.

4 . RESULTS

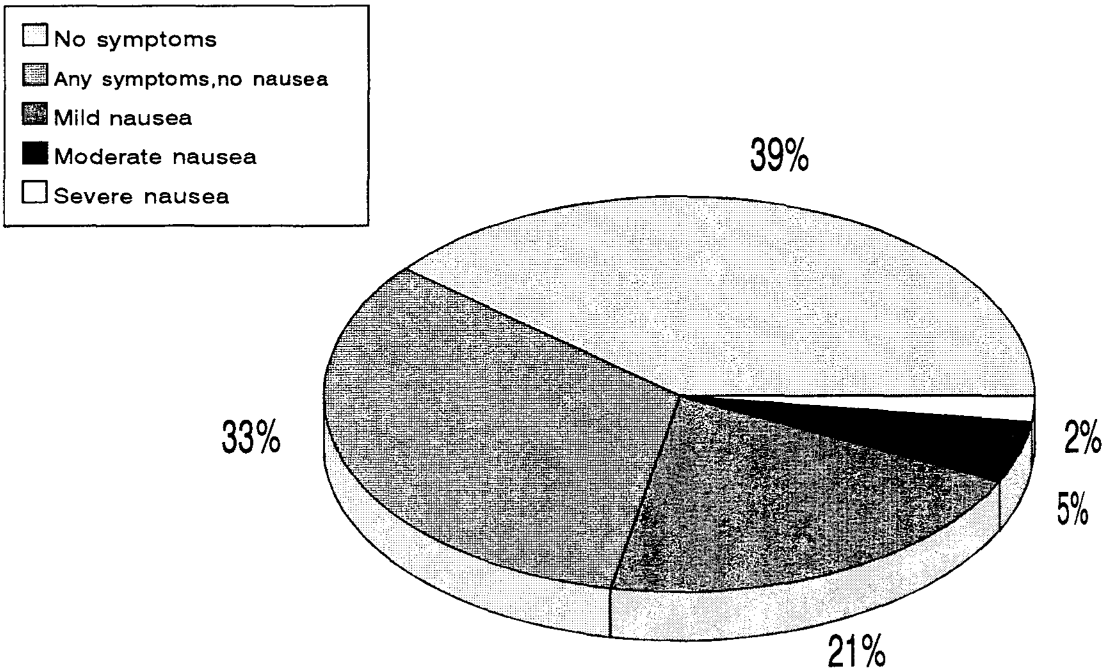
Of the 150 subjects, 4 were excluded from the analyses. These subjects had reported some symptoms at the pre-immersion rating on the malaise scale, and were thus not regarded as being in their normal state of health. Consequently the analysis of the data was carried out for 146 subjects.

4.1 Malaise scale results

Eight subjects withdrew from the experiment (4 of these were civilian subjects and 4 were firefighters). These subjects withdrew due either to severe nausea or severe dizziness. For the purposes of the analysis, those subjects who did withdraw during the experiment were given subsequent immersion ratings on the malaise scale which were equal to the rating on which they withdrew. This was felt to be a conservative approach to predicting the missing immersion ratings, given the likelihood of these subjects reporting increasingly higher ratings had they not withdrawn. The post-immersion ratings (5 minutes post-immersion and 10 minutes post-immersion) were scored as missing for these subjects.

The pie chart below illustrates the percentage of subjects reporting each of the ratings on the malaise scale as their highest across the full immersion and post-immersion period.

Figure 2 : Percentage of subjects reporting each of the malaise scale ratings as their highest



As can be seen from this graph, 61% of the subjects reported ratings greater than 1 as their highest at some stage in the study. Only 39% reported a rating of 1 throughout.

The first bar chart below illustrates the frequency of occurrence of each of the ratings on the 1-6 malaise scale at each of the seven time periods. The ratings are illustrated on the legend for the graph. Ratings of 2 were most commonly associated with dizziness, stomach awareness, headaches, eyestrain, and lightheadedness. The second bar chart presents the data with all the ratings of greater than 1 combined. This illustrates more clearly the progression of symptoms with increasing immersion time.

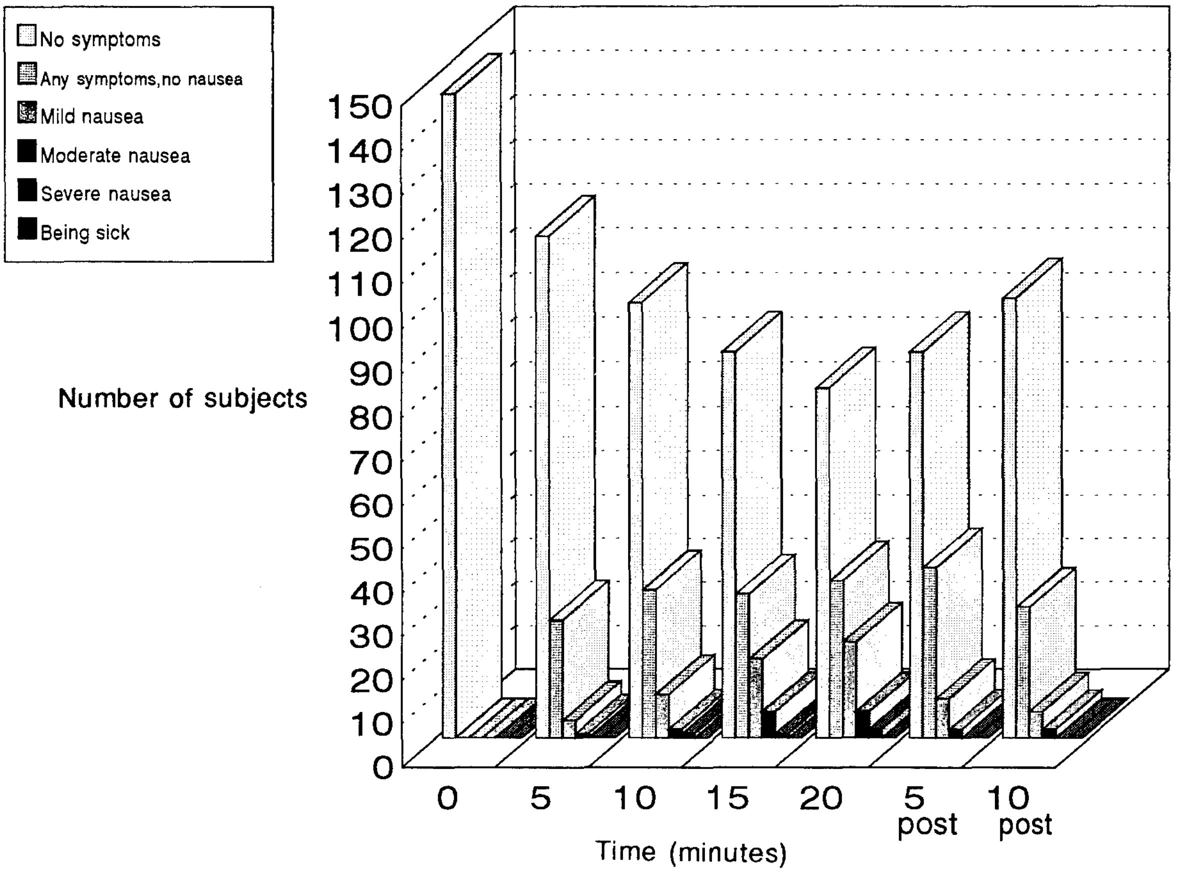


Figure 3 : Frequency of ratings on the malaise scale for all subjects

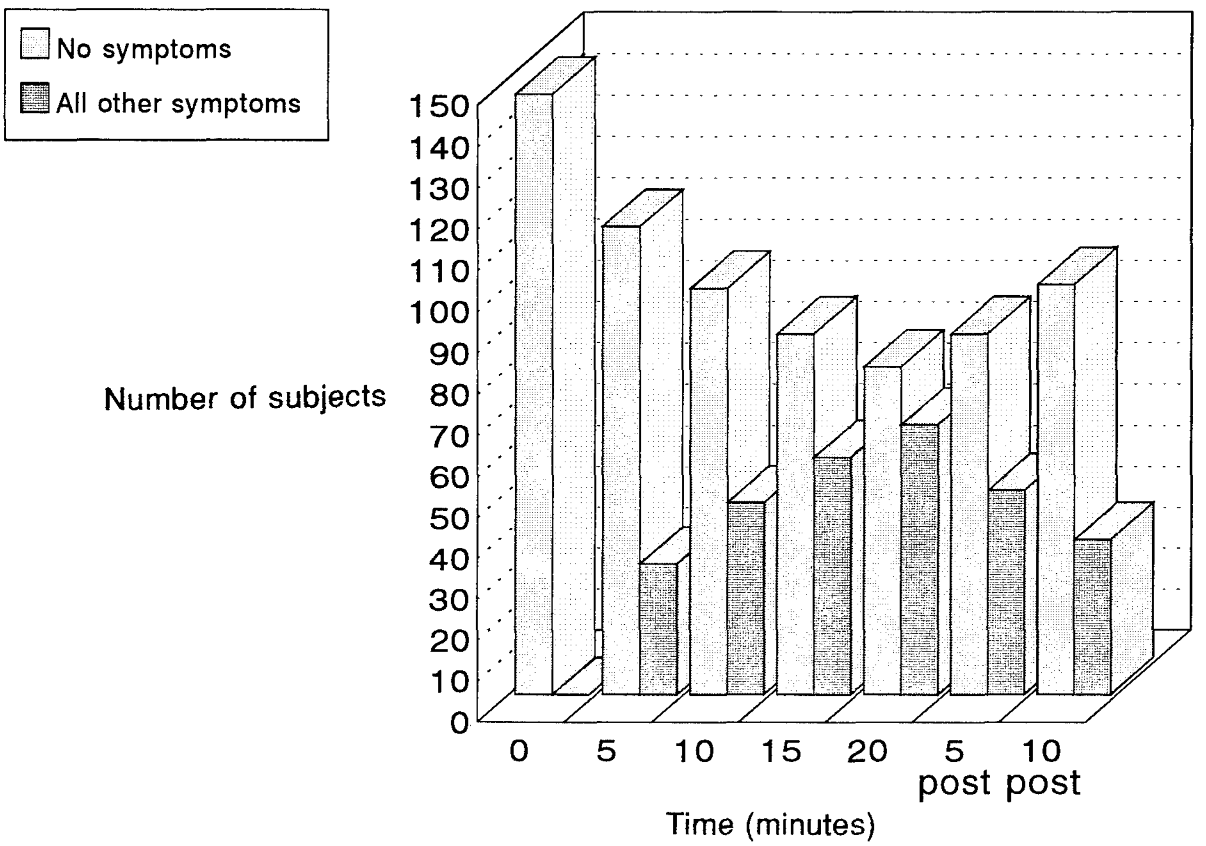


Figure 4 : Frequency of ratings on the malaise scale for all subjects (ratings of 1 and greater than 1 illustrated)

The pattern of reported symptoms across the time periods was very similar for the three groups of subjects - civilian, military and firefighters.

A series of analyses of variance were carried out on the data for the groups of subjects in order to see if there were any significant differences between the three groups in terms of ratings on the malaise scale across the immersion and post-immersion periods. These all yielded non-significant results suggesting no significant differences between the three groups of subjects.

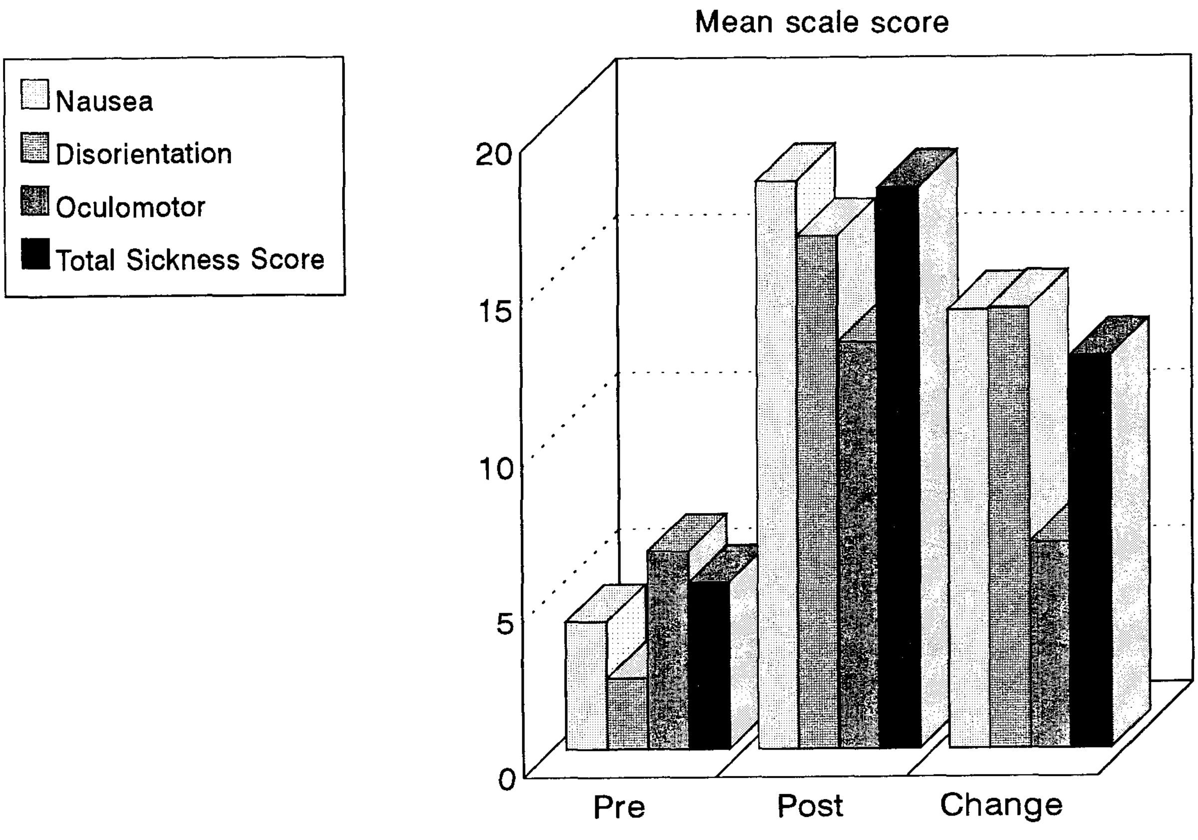
4.2 Pre-immersion and post-immersion symptom checklist results

Following the standard procedure of Kennedy et al (1993) (2), 16 of the questions on the pre-immersion and post-immersion symptom checklists were scored. Scoring the data in this way yields three subscales - Nausea, Oculomotor and Disorientation - and a Total Severity measure.

According to Kennedy, scores on the Nausea subscale are based on the report of symptoms which relate to gastro-intestinal distress such as nausea, stomach awareness, salivation and burping. Scores on the Oculomotor subscale are based on the report of symptoms such as eyestrain, difficulty focusing, blurred vision and headaches. Scores of the Disorientation subscale are related to vestibular disarrangement such as dizziness and vertigo.

The standard procedure is to consider the post exposure profile because of the assumed poor reliability of the difference/change scores that would result from analysis of both pre and post exposure data. However, in view of the large number (54%) of the 146 subjects in this study that reported symptoms on the pre-immersion symptom checklist (this is much higher than other studies of simulator sickness), change score profiles as well as post score profiles were produced. These profiles are illustrated below.

Figure 5 : Pre-immersion, post-immersion and change profiles for all subjects



The profiles were very similar for the three groups of subjects in terms of the pattern and magnitude across the subscales.

These profiles suggest that, with the PROVISION 200 system, nausea is the most significant problem, followed by disorientation and then oculomotor problems.

5. DISCUSSION

This data would enable the side-effects of the VR system used in this study to be compared with the side-effects of other VR systems and other simulators assessed using the same material. The results from this study suggest a high incidence of self-reported malaise resulting from the use of the immersion VR system. The incidence did not differ in a statistically significant manner for the three groups of subjects employed in this study, which suggests that the results found in this study can be generalized to most subject populations. However, there are two reasons why the data must be treated with a degree of caution. Firstly, it must be stressed that the level of symptoms reported in other VR systems may be higher or lower than the level reported in this system. The data presented here can only be cited with reference to the PROVISION 200 VR system with the peripherals used in this study, although the data may be suggestive of a more general incidence of malaise likely to occur with the use of all current VR systems. Secondly, the present experimental procedure could be viewed as having encouraged subjects to dwell on their internal states, and, through constant prompting, to report symptoms that might otherwise have gone unobserved (it is interesting to note in this context that 54% of the subjects reported symptoms on the pre-immersion symptom checklist).

Notwithstanding these reservations, however, it would appear that adverse side-effects are sufficiently common to threaten the success of further studies using the VR system and of applications for the technology in its present state of development - 61% of the subjects in the present sample reported some symptoms of malaise which ranged from symptoms such as headaches and eyestrain to severe nausea; 5% of the subjects had to withdraw from the experiment due to severe nausea or severe dizziness. Consequently some further research has been conducted which has attempted to identify those factors that play a causative role in the side-effects of the VR system. This will be discussed next. This will be followed by a discussion of some areas for future research.

5.1 Further research

Interaction with the environment

Whilst every attempt was made in the experiment detailed above to ensure that all subjects underwent similar experiences during their immersion, differences between subjects in terms of behaviour whilst in the virtual world inevitably occurred. Subjects were free to control their head movements and their speed of interaction with the system. Consequently some subjects clearly moved more slowly and cautiously through the virtual world than others and made fewer head movements. These 'cautious' subjects frequently reported that they would have felt more nauseous had they engaged in more rapid movements. Thus they appeared to have developed coping strategies which enabled them to tolerate the system for the given time period. Interestingly it has been reported that pilots may develop strategies such as restricting head movements to reduce simulator sickness symptomatology.

The extent to which encouraging pronounced head movements and rapid interaction with the system may produce increased nausea was investigated by requiring a new group of 44 subjects who did not take part in the previous study to undergo a fixed set of actions whilst in the VR system. This set of actions was designed to maximise head movements and speed of interaction with the system. The immersion lasted for 10 minutes. The malaise scale ratings of these subjects were compared with the malaise scale ratings of the subjects in the previous study (up to the 10 minute period) who were free to control their head movements and speed of interaction with the system.

However, analyses of variance on the data for the two groups yielded non-significant results at the 5% level (even when the subjects scoring 1 (no symptoms) throughout were removed from the analyses). However the analysis at 5 minutes was significant at the 10% level.

It would appear therefore that no statistically significant difference in malaise scale ratings occurs at the 10 minute immersion point between subjects who are engaging in pronounced head movements and rapid interaction with the system and subjects who are making head movements at their will and moving at their own speed. At 5 minutes some differences may exist. Thus whilst pronounced head movements and rapid interaction may initially cause higher levels of symptoms, subjects appear to adapt to these requirements at 10 minutes. It would consequently appear that factors other than subjects' method of interacting with the virtual environment must be largely responsible for the level of side-effects reported.

Sitting versus standing

28% of the subjects in the study detailed above were found to experience mild, moderate, or severe nausea. Such nausea is a classic symptom of simulator sickness, which is frequently acknowledged as having much in common with motion sickness. We are quite accustomed to motion sickness inducing situations (ie. situations in which information presented to the visual and vestibular systems is contradictory) when seated (eg. in cars and trains) but not when standing up.

In addition, when standing up and interacting with the virtual environment subjects have the facility to make natural walking movements (within a restricted area) as well as movements via the 3D mouse. Some people in the previous study took advantage of this facility and frequently made small physical movements. Such movements may act as a source of confusion when made in conjunction with movements via the 3D mouse and may partly contribute to subjects' experience of adverse side-effects such as dizziness and disorientation. On the other hand, however, such movements may provide useful kinaesthetic and vestibular cues to body position and movement which may attenuate such adverse side-effects.

Consequently it is possible that seating subjects during a VR immersion may affect levels of reported malaise.

44 subjects were immersed in the VR system for 10 minutes. 24 subjects stood during their immersion and 20 subjects were seated.

However, analyses on the data for the two groups yielded non-significant results (even when those scoring 1 (no symptoms) throughout were removed from the analyses) suggesting no significant difference between the two groups of subjects in terms of their ratings on the malaise scale across the immersion period.

It would appear, therefore, that no statistically significant difference in malaise scale ratings occurs between subjects who are standing during their immersion and subjects who are sitting during their immersion.

Inter-pupillary distance

33% of the subjects who experienced symptoms in the initial study reported ocular associated problems - these were eyestrain, difficulty focusing, blurred vision, headaches and visual fatigue.

In the PROVISION 200 system the two LCDs in the headset are a fixed distance apart, with the difference between the images projected to these LCDs set in the software to the average male inter-pupillary distance. The experimental hypothesis was that the subjects who did report ocular problems in the previous study would be those with the greatest inter-pupillary distance deviations from the fixed system configuration. The inter-pupillary distance of 50 of the subjects in the initial experiment was measured using an inter-pupillary distance ruler.

The hypothesis was not found to be supported for the group of 50 subjects as a whole. The only significant finding was related to subjects with an inter-pupillary distance less than the system configuration (which was the majority of subjects). For these subjects there was some suggestion that those subjects with ocular problems did have the greatest deviations from the system configuration.

5.2 Future research

Movement in the virtual environment

One of the areas for future research concerns subjects' method of movement through the virtual environment.

It is likely that the method of movement in the virtual world, via the 3D mouse, makes a significant contribution to the level of reported nausea. It produces a classic motion sickness type situation in which the inputs of the visual, vestibular and kinaesthetic systems are incongruent with each other and previous experience - the visual system suggesting body movement and the other systems suggesting a static body position. The contribution of the method of moving through the virtual world to the reported nausea could be investigated by facilitating more natural methods of movement through the virtual world. One possibility would be to couple subjects' movements on a treadmill to their movements through the virtual world. This would allow subjects to actually walk through a virtual environment thus providing them with all the normal vestibular cues to movement. Levels of nausea would be expected to fall in such a situation.

Habituation

A second area for future research concerns the issue of habituation to the side-effects of immersion in VR. Immersion in VR is an unusual and novel experience. It takes some time to become accustomed to wearing the VR Flight Helmet and to the methods of movement and interaction with the virtual environment. It is possible that repeated immersions in the VR system will produce a decrease in side-effects as subjects become more accustomed to, and confident about, interaction with the system. Clearly there would also be the possibility of a systematic desensitization occurring with repeated exposure. Further research could address the issue of whether subjects will habituate with repeated immersions in VR. There is some suggestion that habituation may lead to reduced symptoms during immersion, but greater levels of post-immersion symptoms. After effects of simulator exposure have been observed in experienced simulator users. Consequently it would be appropriate for research investigating habituation effects to assess levels of malaise amongst subjects over longer post-immersion periods than those employed in the first study reported.

Levels of concentration

Finally, some evidence appears to suggest that concentration levels are related to severity of simulator sickness, with relatively greater degrees of concentration being associated with relatively lower levels of sickness. Some subjects appeared to have to concentrate more than others during the immersion, particularly when using the 3D mouse to pick up and manipulate objects. The effect of varying levels of concentration on adverse side-effects of the system could be experimentally investigated. Clearly if increasing concentration does reduce the severity of side-effects of the system then any further uses of the technology for experimental research purposes or for particular applications should attempt to maximize the concentration levels of the users.

Research into these issues would provide further information on the side-effects of immersion in VR reported, and may help in the identification of methods of reducing these side-effects.

6. CONCLUSION

In conclusion, the main study described in this paper set out to investigate the frequency of occurrence and severity of side-effects of using an immersion VR system. The results of this study suggested that adverse side-effects are sufficiently common to threaten the success of further research using VR and of applications for the technology in its present state of development. Some further research has consequently been conducted which attempted to identify those factors that play a causative role in the side-effects of the VR system. This research and areas for future research have been discussed.

7. REFERENCES

1. Havron M D & Butler L F, 1957: Evaluation of training effectiveness of the 2FH2 helicopter flight trainer research tool. Naval Training Device Center, Port Washington, New York, NAVTRADEVCEN 1915-00-1.

2. Kennedy R S, Lane N E, Berbaum K S & Lilienthal M G A, 1993: Simulator sickness questionnaire (SSQ): a new method for quantifying simulator sickness. International Journal of Aviation Psychology, 3 (in press).

VISUAL ACCOMMODATION TO VIRTUAL IMAGE DISPLAYS WHEN KNOWLEDGE OF OBJECT DISTANCE CONFLICTS WITH OPTICAL DISTANCE

by

John C. Kotulak, Stephen E. Morse and **Roger W. Wiley**
US Army Aeromedical Research Laboratory
PO Box 577
Fort Rucker
Alabama 36362-0577
United States

SUMMARY

In virtual image displays, the image is typically at or near optical infinity, while the object may be at any distance. This can create a conflict between the known distance of a target and its optical distance. If accommodation is drawn to the known distance of the object rather than the optical distance of its image, considerable retinal image blur can result. To determine whether this actually occurs, we measured the accommodation of seven young adult subjects with a dynamic infrared optometer. The subjects viewed a collimated virtual image of a target monocularly through third generation night vision goggles (ANVIS). Although the target itself was positioned randomly at either 6.0, 1.0, 0.5, or 0.33 m from the observer, its image was maintained at infinity by compensatory adjustments of the ANVIS objective lens. The observer was aware fully of the actual distance of the target. A simulated clear starlight night sky condition was used in order to degrade image quality such that the accommodative feedback loop was "semiopen," an intermediate state between the closed and open loop conditions of previous experiments. The results show that for some subjects, knowledge of object distance is a more powerful cue for accommodation than the image's optical distance; however, for the majority of subjects, this is not the case. The subjects who were susceptible to the knowledge of object distance cue reported severe blur when the object was nearby. We also found that these same subjects, i.e., the susceptible ones, tend to have a more proximal dark focus than those whose accommodation is not influenced by knowledge of object distance. The linkage between dark focus and susceptibility to proximal influences has not been previously demonstrated and needs to be explored further.

INTRODUCTION

During ordinary viewing, the observer typically sees a real object, and a clear image of this object is formed on the observer's retinas. The process of changing the focus of the eyes such that the retinal images remain clear at varying viewing distances is called "accommodation." Accommodation is guided by both physiologic and psychologic stimuli. Retinal image blur is the principal physiologic stimulus (1), and perceived object distance is the principal psychologic stimulus (2). Normally, retinal image blur and perceived distance act in harmony in that when both cues are available, accommodation is more accurate than when only one of them is present (3).

However, under virtual reality conditions such as in flight simulators or in aircraft equipped with helmet mounted displays, a mismatch can occur between retinal image blur and perceived distance. This is because the observer no longer sees the real world, but instead views an optical image of the world. Although the image typically is placed at or near optical infinity, it may convey a psychological sense of nearness. This sense of nearness may derive from the object which is being imaged, if the object is something that the observer would expect to find close by, such as the flight controls in a simulator (4). The sense of nearness also could derive from the close proximity of the display to the eye.

Under optimal conditions accommodation can deal effectively with the mismatch between conflicting physiologic and psychologic cues (5-7). Such conditions are referred to as "closed loop," which describes the state of the negative feedback loop of the accommodative control system when target contrast and luminance are high, and when the quality of the

Presented at an AGARD Meeting on 'Virtual Interfaces: Research and Applications', October 1993.

retinal image is not degraded otherwise. Under closed loop conditions, the physiologic cues predominate over perceived nearness and the retinal image remains clear when there is a conflict between cues (5-7).

When cues from retinal image blur are removed, such as by increasing the depth of focus of the eye by viewing through a small artificial pupil, the situation is quite different. When this happens, the perceived nearness of the target is highly influential in determining the level of accommodation (3, 8, 9). Such conditions are referred to as "open loop" because negative feedback information about retinal image blur is denied to the accommodative control system.

During the viewing of virtual reality displays, however, the accommodative loop is probably neither completely closed nor open, but rather "semiopen." The semiopen loop state is the result of the limited spatial resolution that is found in such displays, and perhaps due to reduced luminance and contrast, and the presence of dynamic visual noise. These characteristics result in decreased accommodative accuracy (10), presumably because they make it difficult for the visual system to detect retinal image blur, and thus respond to it by changing accommodation.

Our purpose in the present study was to determine the extent to which accommodation is influenced by psychological factors under viewing conditions similar to those found in virtual reality systems. To do so, we performed an experiment in which accommodation was measured during viewing through an optical instrument which creates the semiopen loop condition that is typical of virtual reality displays. In this experiment, we created a conflict between cues from retinal image blur and perceived distance by varying target distance over a wide range, while holding the image of the target constant at optical infinity.

METHODS

The optical instrument was a pair of night vision goggles (ANVIS), which are unity magnification devices that electronically amplify ambient light and provide a photopic visual display under night sky conditions. The night vision goggle image creates the semiopen loop condition which we desired for this experiment because of its relatively low luminance (1 cd/m^2), its low spatial frequency content (the -3 dB rolloff of the spatial modulation transfer function is at 5 cycles/degree), and the presence of uncorrelated dynamic visual noise (11). The night vision goggle display luminance was achieved by adjusting the ambient luminance to the level of clear starlight.

The visual stimulus was a Bailey-Lovie visual acuity chart. Due to its design, this chart provides targets of the same visual angle at each test distance that was used in the experiment (6, 1, 0.5, and 0.33 m). The Weber contrast of the letters on the chart, when viewed through the night vision goggles, was 65 percent.

Accommodation was measured monocularly under steady-state conditions with a dynamic infrared optometer. The steady-state values were calculated from the mean of 600 samples (20 samples/sec X 30 sec/trial). In addition to measuring accommodation during instrument viewing, we also measured accommodation in complete darkness. The so-called dark focus of accommodation is the resting point of the accommodative control system (12).

Object distance was varied randomly over the test range, while image distance, size, luminance, and contrast were held constant. The instrument eyepieces were set to 0.0 D and the objective lenses were focused for the object distance. The subject was informed of object distance and was instructed to observe as the test distance was measured out. The subject's task was to view through the instrument and keep threshold-sized letters clear. Seven young adult volunteer subjects were used. All subjects were either 20/20 or corrected to 20/20 for the target distance, and were free from eye disease or other ocular anomalies.

RESULTS

Figure 1, in which each plot represents a different subject, shows how instrument accom-

modation varied with object distance. Negative values of accommodation represent accommodation which is less than that required to fully compensate for a hyperopic refractive error. The subjects seem to fall into two distinct groups, that is, those affected by changes in object distance (n = 2), and those unaffected (n = 5). The affected subjects readily perceived target blur at the nearer object distances, but reported that they were unable to

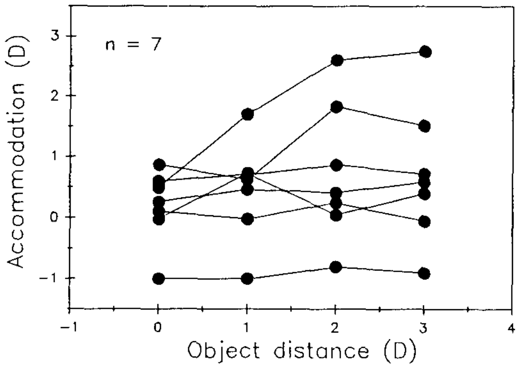


Figure 1. Accommodation as a function of object distance when the data of each subject are shown individually.

eliminate the blur. In Fig. 2, the responses of the subjects within each group are averaged, and the mean dark focus of each group is shown. The error bars indicate one standard deviation. The dotted line with arrow indicates the mean dark focus of the susceptible group, while the solid line with arrow indicates the mean dark focus of the nonsusceptible group. Thus, the group with the more proximal dark focus is the one that was affected by changes in object distance.

The subject who exhibited the most susceptibility to the effect of object distance was retested on a subsequent day. There was no statistically significant difference in instrument accommodation for this subject between the 2 days ($t = 1.23$, $p = 0.31$). In addition, the dark focus of each subject was measured immediately pre- and posttest. There was no evidence of a change in dark focus ($t = 0.33$, $p = 0.75$).

DISCUSSION

Our results indicate that knowledge of object distance can influence the level of accommodation under semiopen loop conditions. This

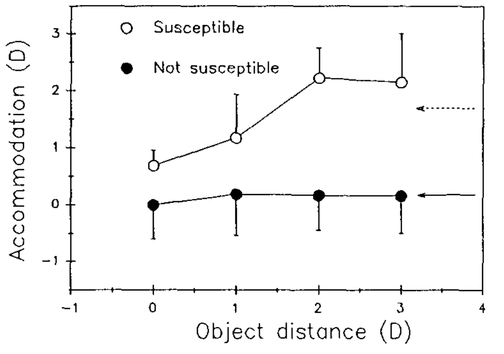


Figure 2. Accommodation as a function of object distance when the subjects are grouped according to susceptibility to proximal cues.

is predictable from earlier works which showed no proximal effect for closed loop conditions, but a pronounced effect for open loop conditions. Perhaps more significant is that the proximal effect appears to be all or none, rather than graded. This is not predictable from previous studies, and neither is the apparent relationship between susceptibility to the proximal effect and dark focus magnitude. Current theory does not explain why individuals with proximal dark focuses should be more susceptible to perceived nearness than individuals with distal dark focuses.

Caution must be used in extrapolating from the results of the present experiment to most existing virtual reality systems. This is because virtual reality displays are typically binocular, and our experiment was done under monocular conditions. Under binocular conditions, accommodation tends to be more accurate, and probably less susceptible to psychological influences, than under monocular conditions. This is due to "vergence accommodation," which is present under binocular but not under monocular conditions. However, the effects of vergence accommodation vary among subjects, so that subjects in whom vergence accommodation plays little or no role may be influenced by perceived nearness even during binocular viewing.

REFERENCES

1. Campbell, F. W., and Westheimer, G., "Factors Influencing Accommodation Re-

sponses of the Human Eye," J. Opt. Soc. Am., 48, 1960, pp 568-571.

2. Ittelson, W.H., and Ames, A., "Accommodation, Convergence, and their Relationship to Apparent Distance," J. Psych., 30, 1959, pp 43-62.

3. Kruger, P.B., and Pola, J., "Dioptric and Non-dioptric Stimuli for Accommodation: Target Size Alone and with Blur and Chromatic Aberration," Vision Res., 27, 1987, pp 555-567.

4. McCauley, M. E., and Sharkey, T. J., "Spatial Orientation and Dynamics in Virtual Reality Systems: Lessons from Flight Simulation," in "Proceedings of the Human Factors Society 35th Annual Meeting," 1991, pp 1348-1352.

5. Hennessy, R. T., "Instrument Myopia," J. Opt. Soc. Am., 65, 1975, pp 1114-1120.

6. Morgan, M. W., "Accommodation and Convergence," Am. J. Optom. Arch. Am. Acad. Optom., 7, 1968, pp 417-454.

7. Alpern, M., "Vergence and Accommodation: Can Change in Size Induce Vergence Eye Movements?" Arch. Ophthalmol., 60, 1958, pp 355-357.

8. Rosenfield, M., and Gilmartin, B., "Effect of Target Proximity on the Open-loop Accommodative Response," Optom. Vision Sci., 67, 1990, pp 74-79.

9. McLin, L. N., Schor, C. M., and Kruger, P. B., "Changing Size (Looming) as a Stimulus to Accommodation and Vergence," Vision Res., 28, 1988, pp 883-898.

10. Kotulak, J. C., and Schor, C. M., ""The Effects of Optical Vergence, Contrast, and Luminance on the Accommodative Response to Spatially Bandpass Filtered Targets," Vision Res. 27, 1987, pp 1797-1806.

11. Riegler, J. T., Whiteley, J. D., Task, H. L., and Schueren, J., "The Effects of Signal-to-noise Ratio on Visual Acuity through Night Vision Goggles," U.S. Air Force Armstrong Laboratory Technical Report No. AL-TR-1991-0011, 1991.

12. Leibowitz, H. W., and Owens, D. A., "Anomalous Myopias and the Intermediate Dark Focus of Accommodation," Science, 189, 1975, pp 646-648.

Human use: Human subjects participated in these studies after giving their free and informed voluntary consent. Investigators adhered to AR 70-25 and USAMRDC Reg 70-25 on Use of Volunteers in Research.

Low Cost Software-Based Rendering and Stereoscopic Interfaces for Teleoperation and Virtual Reality

Dave M. Stampe
SR Research
55 Oakmount Rd. Ste. 1108
Toronto, Canada M6P 2M5
dstampe@psych.toronto.edu

Julius J. Grodski
DCIEM, Toronto, Canada

SUMMARY

Many interface systems require generation of 3D graphics, whether as the entire display in virtual reality systems or as an overlay on live video in teleoperation. Costs must be kept low to make such systems practical, but real-time response speed must not be sacrificed.

Described here is a very low-cost rendering and VR support package for 386 and 486 PCs, which requires no added hardware to achieve real-time drawing rates (20 to 60 frames per second). It includes integral support for generation and display of stereoscopic graphics in many formats, including field-alternate displays using LCD shutter glasses, and wide-angle head-mounted displays. Many common PC interface devices are supported, including mouse, joystick, 6D pointing devices, and head trackers.

Inexpensive PC multimedia cards allow output to be recorded on a VCR, or overlaid onto live video, including stereoscopic TV images from teleoperated remote cameras. Full source code is available, allowing the software to be customized for any application.

1. Introduction

Generation of three-dimensional computer graphics is a basic requirement of many computer interfaces, from CAD design tools to scientific visualization and virtual reality (VR). Such graphics range from simple wireframe drawings to photorealistic raytraced images used in computer art and movies. The most demanding of all computer graphics applications are those that require images to be produced in real time, such as flight simulators and VR. Expensive special-purpose hardware is often needed to achieve required drawing speeds of 10 to 60 frames per second.

In applications where depth judgments are important, stereoscopic imaging of live video or computer-generated graphics is advantageous. It requires the generation of suitable left and right eye images from two video cameras or by the computer such that when viewed by an observer, they simulate the process of left/right "live" viewing of the environment. The observer interprets differences between the left and right eye images as depth information in the scene.

Stereoscopic imaging using video has been in use for decades in the field of teleoperation and telerobotics. Real-time stereoscopic computer graphics is a capability that is much more recent then stereoscopic video, as twice the drawing rate or a second set of rendering hardware is required to produce both eye's views. Stereoscopic computer graphics are an essential part of Virtual Reality (VR) applications, while in teleoperation it is advantageous to have the capability to overlay stereoscopic computer graphics on stereoscopic video. Computer-generated imagery has seen limited use in teleoperation due to the cost and limited power of graphics generators and workstations capable of producing real-time three-dimensional graphics. Stereoscopic graphics are rarely used, as they are expensive in terms of computer resources and require careful attention to real and "virtual" (computer graphics simulated) camera setup.

2. Augmented Teleoperation

Early uses of stereoscopic computer graphics in teleoperation focused on the simulation of a manipulator operating in a remote task environment [1,2,3]. Fully immersive systems using head-mounted displays (HMD), such as the Virtual Reality Systems [4] and the Virtual Interface Environment Workstation (VIEW) [5,6] were state of the art as of the mid-late 80's, and inspired the development of totally computer-generated virtual environments, as in today's VR systems.

Concurrently, low-budget methods involving the combination of remote stereoscopic video and real-time computer graphics were evolving [7,8]. The graphics were generated by a Commodore Amiga computer, which used a M68000 processor and included custom graphics hardware to speed wireframe rendering. A genlock overlay device allowed display of the graphics overlaid on stereoscopic video on the computer monitor, and LCD glasses were used to view the image. Video was generated remotely from two color CCD cameras, and could be recorded for later analysis.

Although the Amiga was the fastest low-budget solution for real-time stereoscopic graphics available at the time, more imaging power was needed to render complex moving images. Advanced hardware such as the Silicon Graphics IRIS graphics workstations have been used more recently in the work at the Jet Propulsion Laboratory [9] and at the University of Toronto [10,11], but at a substantial increase in system cost. One advantage of the IRIS for graphics generation is that extensive graphics support libraries are available, whereas the Amiga software had to be written by hand.

Presented at an AGARD Meeting on 'Virtual Interfaces: Research and Applications', October 1993.

3. Low-Cost Rendering Systems

The most common personal computer in use today is the IBM PC family: there are almost 10 times as many IBM PC compatible computers in use as any other design. The newer designs based on Intel i386 and i486 processors (and the latest Pentium processor) have computational power exceeding many low-end workstations. There is a very large base of programming talent available for these machines, and therefore this would be the ideal platform for development of graphics systems. Until recently, there was little general-purpose three-dimensional rendering tools and no real-time graphics software available for these machines. The PC is a difficult platform for real-time graphics due to its hardware design, and the techniques needed were proprietary to video game authors and commercial developers.

Ideally, a low-budget rendering system for general purpose use would be based on the 386 or 486 IBM PC, would require little or no special hardware, and would allow substantial modifications to be made by the programmer. This would require the availability of most or all of the source code, and enough information to write support software for new interface and display devices. It should be able to draw images in real time: at least 10 frames per second. Photorealistic rendering and high resolution may be sacrificed to achieve these drawing speeds, but wireframe graphics should be avoided at they decrease 3D depth cues such as interposition. Software support for generation of stereoscopic graphics and drivers for common stereoscopic display devices is essential, as this is one of the most difficult and least-documented aspects of 3D graphics implementation.

4. REND386

A software toolkit for real-time three-dimensional and stereoscopic graphics has been developed, and is available free of charge to programmers. The capabilities and some applictions of REND386 are described below.

4.1 Background

The REND386 project began as part of a worldwide effort by experimenters on the Internet to develop low-cost personal virtual reality systems using widely available and low-cost technology. Building low-cost head-mounted displays and other interface devices turned out to be much easier than generating stereoscopic graphics in real time. The Amiga was considered to be the fastest graphics computer available, but its internal graphics accelerator hardware was unsuited to the speed and detail of 3D graphics required for VR.

The PC has one of the largest hardware and programmer bases of any computer, and was the system of choice for the project. Developing real-time graphics software for the PC requires extensive knowledge of the complex interactions between the PC's VGA display system and the processor, and carefully optimized assembly code is required to implement drawing and rendering kernels. Extensive use of mathematics is required for 3D rendering as well [12], and much of this must be coded in assembler to achieve the needed speeds. The renderer project was undertaken by D. Stampe and B. Roehl [13] and evolved into the REND386 graphics and VR programmer's toolkit. The software is written in assembler and C, and runs on 386 or 486 PCs. A math coprocessor is not required.

4.2 3D Renderer

The core of REND386 is a real-time 3D graphics renderer, written in assembler and C. It consists of a highly optimized software pipeline which performs visibility testing, coordinate transformations, clipping, lighting, and depth sorting of polygons. Drawing of polygons is performed by display-specific software video drivers, which also support display operations such as image clears, copies, and display of text. The standard drivers support the PC-standard VGA card in 320 by 200 pixel resolution, which is ideally suited to the resolution of most head-mounted displays. Some users have written drivers to support higher resolution displays or graphics accelerator boards.

The renderer is implemented entirely in software, and its performance depends on the speed of the computer and on the type of video driver and VGA card used. A very powerful yet inexpensive graphics generation system may be built for under US$2000, using a 66 MHz i486DX2 processor and a local-bus VGA card. The system can achieve rendering speeds of over 15,000 cosine-lit polygons per second, drawing a minimum of 35 frames per second with up to 500 polygons visible on screen. This speed is sufficient to allow generation of both left and right eye images on the same PC at speeds sufficient for real-time stereoscopic VR.

Speeds are dependent on how many objects are visible on screen, as objects that are not visible are eliminated early in the rendering pipeline. The pipeline has many visibility-based optimizations, typically pruning a 3000 polygon virtual world to less than 250 polygons before reaching the time-intensive lighting and drawing stages. Polygons may be rendered in metallic and pseudo-glass styles as well as solid or lit colors. Lines and wireframe rendering are also supported.

The images are rendered as seen from a virtual viewpoint, specified by either position and Euler angles or by a homogenous matrix. The renderer is controlled by a viewport description containing viewpoint data as well as the size and position of the window on the display into which the image is to be drawn, and the field of view of the display. The field of view is used to set the correct degree of perspective to match the display device: a small desktop monitor image may have a field of view of 15°, while a head-mounted display image may cover in excess of 120°. The renderer is also capable of offsetting the center of the image in the view window, required to adjust apparent depth in stereoscopic viewing. It can also render images that are horizontally or vertically inverted to match display devices that use mirrors. The renderer is capable of displaying a two-color horizon background, and a three axis "compass" to help orient the viewer during exploration.

Lighting of objects is very important to 3D rendering and VR, as it increases the apparent depth of objects and

prevents masking of objects against similarly-colored backgrounds. The renderer supports two point or directional light sources and one diffuse light source for illumination of objects. Lighting is computed as independent of distance and is proportional to the cosine of the angle between the light source and each polygonal facet of objects. Each polygon has a reflectance and hue value that are combined with the computed lighting strength to determine which of the 256 available colors in the VGA palette will be used to draw the polygon.

4.3 Stereoscopic Support

One of the most important tools available in REND386 is integrated stereoscopic imaging support. Given information about the display such as screen size and the distance from the viewer, it will compute the proper field of view, left and right viewpoint positions, and offsets of images in the view window to create an orthoscopic stereoscopic display. Orthoscopic displays show objects in the world at proper depths relative to the viewer, as if the monitor were a window into the virtual world, but may cause eyestrain when viewed on small monitors. Non-orthoscopic views are often required, for example to exaggerate the sense of depth or to make objects float in space in front of the monitor. These may be achieved by modifying the stereoscopic model that REND386 uses to compute the view parameters.

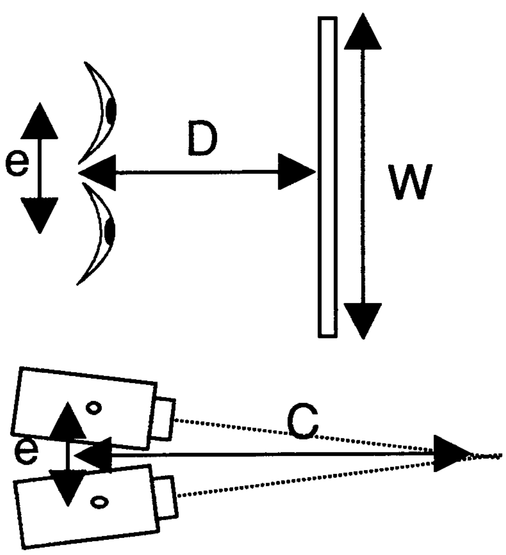


Figure 1. The REND386 stereoscopic model computes imaging parameters based on physical dimensions: e=eye spacing, D=screen distance, W=screen width, C=convergence distance. These are used to compute perspective (field of view), left and right viewpoint coordinates in virtual world, and offset of left and right images on screen.

The stereoscopic calculations used by REND386 are based on the camera/monitor system often used in teleoperation, shown in Figure 1. The two cameras are spaced by the same distance as the viewer's eyes, and are pointed so their optical axes converge at a known distance, usually the same as the distance from the viewer to the monitor screen. The images from the cameras are displayed on the monitor alternately, with LCD shutter glasses used to ensure the images reach the proper eyes. The parameters of the system are then eye and camera spacing, convergence distance, stereoscopic window size, and viewer distance. Internal calculations also require the world scaling factor, in order to relate the arbitrary numerical scale of the virtual world to the real world. A scale of 1.0 unit to 1.0 mm is often chosen.

The calculation of viewport parameters from these factors is documented in [14]. Exaggerated depth is obtained by increasing the eye spacing parameter, which in combination with changes in world scale can make objects appear miniaturized and close to the viewer. Objects can also be brought out of the monitor by increasing the convergence distance. A wider field of view and exaggerated perspective may be achieved by decreasing the screen distance parameter. All stereoscopic model parameters can be changed through the renderer configuration file without recompiling the code. Fine tuning of these parameters is often needed to suit different viewers, and may be done interactively from the keyboard.

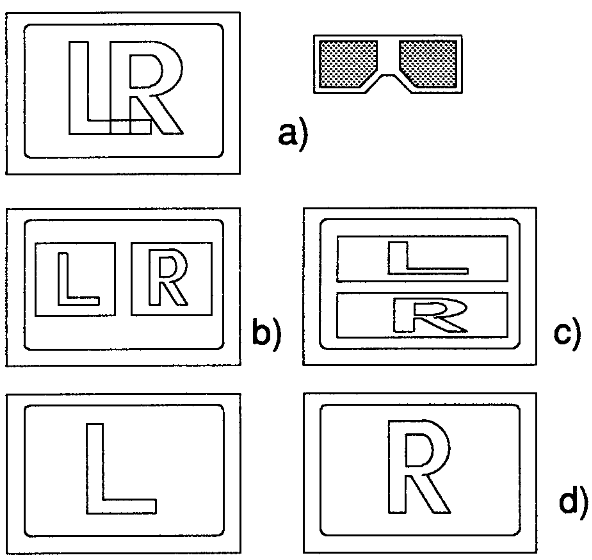


Figure 2. Some of the stereoscopic display modes supported by REND386: a) Time-multiplexed stereo with LCD shutter glasses. b) Side-by-side stereo windows for stereopticon viewers. c) Vertical stereo windows for double-speed viewers such as Stereographics CrystalEyes. d) Seperate VGA displays for head-mounted displays.

4.4 Stereoscopic Display Devices

Time-multiplexed stereo is directly supported by REND386: all that is required is to connect a pair of LCD glasses to one of the computer's serial ports through an inexpensive driver circuit and to enable the stereoscopic display. REND386 supports other types of stereoscopic displays in addition to the time-multiplexed monitor display, some of which are illustrated in Figure 2. Two windows for left and right eye images may be displayed, for

use with lens or mirror viewing devices. The windows may also be stacked vertically for use with scan-doubling viewers such as the Stereographics CrystalEyes system. The renderer also supports VGA display cards with independent video buffers, for generation of independent left and right eye views for head-mounted displays. Special display control parameters are available to support these devices, all of which may be controlled through the configuration file.

The stereoscopic model used by REND386 assumes that the left and right eye images will be displayed in the same window on the screen. If separate screen windows or separate displays in an HMD are to be used, the left and right eye images must be offset horizontally to compensate. This correction is also useful in HMDs to compensate for differences in user's interpupillary spacings. Some optical systems may produce images that are tilted, such as the wide-angle HMD in Figure 3, and require compensatory rotation of the image plane during rendering. Such rotation may be set independently for the left and right eye images.

The choice of the PC as the platform for REND386 has made a number of inexpensive multimedia products available for use with 3D graphics. Inexpensive video-overlay boards such as the VideoBlaster from Creative Labs can combine live video with 3D graphics from the VGA display for augmented teleoperation. With special drivers, these cards can be adapted to overlay stereoscopic graphics from REND386 onto standard field-sequential stereoscopic video images, or to convert multiplexed stereoscopic video into left and right eye images for head-mounted displays. VGA-to-video converters such as the AVerKey from ADDA Technologies convert the VGA images produced by REND386 into video for driving HMDs or for recording of normal and field-sequential stereoscopic images on videotape.

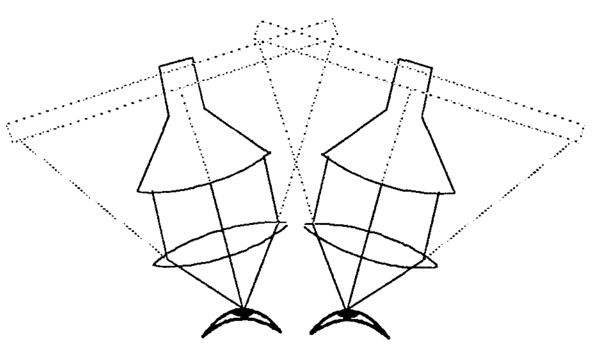


Figure 3. Example of a display which requires imaging with rotated image planes. The angled configuration of lenses allows a much greater peripheral field of view and allows larger display devices to be used than otherwise possible.

4.5 Environments and Simulation

A virtual world consists of a collection of objects loaded into REND386 for viewing or manipulation. Objects consists of polygons, and are loaded from PLG files containing lists of vertices and polygon descriptions. Multiple objects can be loaded and arranged under control of a WLD file, which describes a complete REND386 world. The world can also be split into smaller areas for visibility control and navigation.

For motion control, objects may be connected together by articulated joints into hierarchical figures. For example, athe robot arm in Figure 4 could be created by attaching gripper finger objects to a wrist object, and the wrist to an arm object. If the arm-wrist joint is rotated, the wrist and gripper fingers will move as an indivisible object. Joints may also be moved as well as rotated to change the relative positions of objects. Specifying figure configuration by angles of joints is much more useful than explicitly computing positions of each of the parts, and makes animation or tracking of real-world objects much easier.

The joint mechanism is implemented by cascading homogenous matrices to describe each object position. Each figure is arranged as a tree of joints and objects, which helps to organize motion and allows efficient updating of object positions. In most rendering systems, such hierarchical figures are implemented during rendering by cascading viewing transforms from each joint matrix. REND386 actually moves each object in the world when it is rendered, and caches the new positions of the objects. Because the actual position in the world of all objects is available, collision detection can be performed efficiently.

REND386 contains extensive fixed-point libraries for matrix operations such as inversion and transformation. Trigonometric function are also available, including four-quadrant arctangent, and Euler angle to matrix and matrix to angle translations. The fixed-point formats were designed to achieve near floating-point precision and range, while matching or exceeding the 387 floating-point coprocessor in speed. For example, matrix entries and trigonometric results are accurate to 8 decimal places, and world coordinates have a range of more than 7 significant digits. These libraries and the articulated figure support are essential for tools for creation and manipulations of virtual environments, or for representation of real-world events for augmented reality presentations.

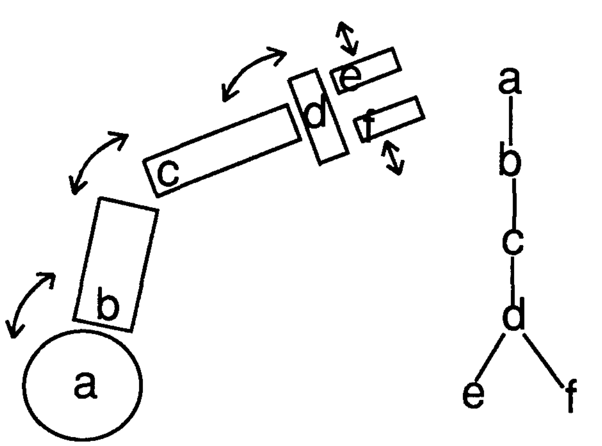


Figure 4. Example of an articulated figure for simulation. Each of the objects in the figure pivots around the lettered location, with respect to the object above it in the hierarchy shown on the right. The "fingers" e+f are translated rather than rotated with respect to the "wrist" d. All joint positions and angles may be set by data from a real-world robot arm.

4.6 Virtual Reality Toolkit

REND386 was developed especially to provide a means for creation of simple virtual reality systems. The renderer easily produces the 10 frames per second required for usable VR systems [15] even when producing both left and right eye images for stereoscopic presentation via HMD. Support for three-dimensional manipulation devices and even an inexpensive gestural interface device (the Nintendo PowerGlove) is built in. Navigation and manipulation devices such as mouse, trackball, and joystick are supported, as are head trackers. New devices can be interfaced by writing loadable drivers or by modifying the software.

Head-mounted displays are supported by modifying the stereoscopic display model to match the field of view and display spacing of each device. There is little standardization in HMD design, and parameters may vary even between eyes of the same display. Special video cards or VGA-to-video converters may be used to drive the HMD displays. The stereoscopic rendering support is designed to be flexible enough to support almost any display device.

REND386 includes a mechanism to integrate head tracker data into viewpoint control for proper image generation to the HMD. The articulated-figure mechanism is used to build a body for the viewer, to which head-viewpoint and hand-manipulators can be attached. The head tracker then controls the head-body joint, and the 3D manipulation device controls the hand-body joint. The body can then be moved through the virtual world, with the hand and viewpoint properly attached. The body can also be attached to moving objects in the world, allowing the viewer to "ride" objects or to be moved in complex trajectories.

Creation and control of autonomous objects in the world can be performed through a simple animation control language that allows objects perform cyclic motions, to react to the presence of the viewer or to be triggered by the manipulation device. More complex motions, such as copying the motions of a real-world device, can be implemented by adding C code to the system to directly control joint positions.

5. Conclusions

Virtual reality, visualization and teleoperation augmented with stereoscopic graphics require real-time three-dimensional rendering of graphics. Low cost is important for many applications, but development of rendering software from scratch is out of the question in most cases.

REND386, a software toolkit for realtime graphics on the IBM PC platform, includes many features to make experimentation and development of new systems easy. Virtual worlds and environments are easily created and modified, and can contain autonomous objects. It is possible to interface the virtual world to the real world by adding object-control code: for example, an articulated model of a robot arm could be moved in synchrony with a real arm by transferring real joint angles to the model. This could be used in an augmented teleoperation application.

Stereoscopic displays are easily created with REND386, and can support a variety of display devices, including LCD glasses and HMDs. Video overlay boards allow stereoscopic graphics and stereoscopic video to be combined for operator aids or comparisons. For example, wireframe outlines of a moving virtual object could superimposed on video of the real object to judge the match between modeled and actual motions.

REND386 supplies a substantially complete software base for 3D graphics and stereoscopic displays for the IBM PC and compatible computers. These computers can provide a cost-effective and widely supported platform for experimentation and applications. Unlike other rendering or graphics packages, full source code is available to the programmer, allowing new display modes or new interface devices to be added.

6. Availability

REND386 is available in the form of a demonstration executable or as source code by FTP from several sources on the Internet. The most recent upgrades are always available from sunee.uwaterloo.ca. Documentation on file formats is included with the demonstration software and is explained in detail in [14]. The software and source code is available free of charge for non-commercial uses.

7. References

(1) T.B. Sheridan, "Supervisory Control of remote manipulators, vehicles and dynamic processes: Experiments in command and display aiding", Advances in Man-Machine Systems Research, Vol 1, 1984, pp.49-137.

(2) K. Fernandez & E. Hinman, "The use of computer graphic simulation in the development of on-orbit telerobotic systems", Robots 11: 17th International Symposium on Industrial Robots, April 1987, 21: 45-53.

(3) W.S.Kim, F. Tendick & L.W.Stark, "Visual enhancements in pick-and-place tasks: Human operators controlling a simulated cylindrical manipulator", IEEE J.Robotics & Automation, Vol. RA-3(5), Oct. 1987, 418-425.

(4) K.B. Evans, "Head Mounted Displays for Telepresence and Workstations", Proc. Second Workshop on Military Robotics Applications, Kingston, Ont.,Canada, 8-11 Aug, 1989, pp. 50-56.

(5) S.S. Fisher, E.M.Wenzel, C. Coler, and M.W. McGreevy, "Virtual interface environment workstations", 32nd Proc. Human Factors Society, 1988, 91-95.

(6) J.D. Foley "Interfaces for advanced computing", Scientific American, Oct.1987, 127-135.

(7) P. Milgram, D. Drascic, and J.J. Grodski, " Stereoscopic Video + Superimposed Computer Stereographics: Applications in Teleoperatrion", Proc. Second Workshop on Military Robotics

Applications, Kingston, Ont., Canada, 8-11 Aug, 1989, pp. 20-25.

(8) P. Milgram, D. Drascic, and J.J. Grodski, "Enhancement of 3-D video displays by means of superimposed stereo-graphics", Proc. 35th Annual Meeting of Human Factors Society, (pp.1457-1461). San Francisco, CA., 1991.

(9) W.S.Kim and A.K.Bejczy, Graphics Displays for Operrator Aid in Telemanipulation", IEEE SMC, 1991, pp.1059-1067.

(10) J.J.Grodski, P. Milgram, and D. Drascic, "Real and Virtual World Stereoscopic Displays for Teleoperation", Proc. 31 DRG Seminar on Robotics in the Battlefield, NATO/DRG, Techn. Proc. AC/243-TP/3, C2.0-C2.16

(11) D. Drascic, J.J. Grodski, P. Milgram, P.Wong, S. Zhai, and K. Ruffo, "SV+SG: Stereoscopic Video with Stereoscopic Graphics", Video Proc. INTERCHI'93, Amsterdam, 25-29 April 1993.

(12) J.D.Foley, A. van Dam, S.K.Feiner, and J.F.Hughes, "Computer Graphics: Principles and Practice", Addison-Wesley. 1990

(13) D.M. Stampe and B.Roehl. "The REND386 Project -- History and Current Status." In *Virtual Realities: Anthology of Industry and Culture*.(In press)

(14) D.M. Stampe, B.Roehl, and J.Eagan. "Virtual Reality Creations", Waite Group Press, Corte Madera CA. 1993

(15) S. Bryson "Implementing Virtual Reality: Virtual Reality Hardware" SIGGRAPH '93 Course Notes #43, pp. 1.3.1-1.3.26. 1993

APPROCHE EXPÉRIMENTALE DU SON 3D : MÉTHODOLOGIE ET RÉSULTATS PRÉLIMINAIRES

L. Pellieux*, C. Gulli**, P. Leroyer*, B. Piedecocq*, A. Leger**, J.P.Menu*

* CENTRE D'ETUDES ET DE RECHERCHES DE MÉDECINE AÉROSPATIALE
BP 73 91223 Bretigny sur orge cedex France

** SEXTANT AVIONIQUE
B.P.91 St Medard en jalles cedex France

SOMMAIRE

Le principe du son 3D, ses domaines d'application et les mécanismes utilisés par l'homme pour localiser un son, sont rappelés. Les travaux réalisés sont décrits.
Le but est la mesure des performances humaines en localisation d'une source sonore simulée à l'aide d'un casque stéréophonique. L'étude peut être décomposée en 2 parties : mesure individuelle des caractéristiques de diffraction acoustique de la tête (fonctions de transfert de tête) de 4 sujets, puis test des performances en localisation, les sujets écoutant des sons élaborés à partir de leurs propres fonctions de transfert. Le sujet doit viser la source perçue, le plus précisément possible, le temps de réponse n'étant pas pris en compte. Les sources ne sont pas matérialisés et le sujet est dans une quasi obscurité. Les résultats montrent que la localisation en gisement est facile pour les 4 sujets. Par contre, la localisation en site est très difficile pour 2 sujets. Pour le meilleur sujet, l'erreur sur la ligne de visée est de 6° en valeur efficace.

INTRODUCTION

A l'origine des travaux présentés, il y a la constatation de la quasi saturation des fonctions visuelles du pilote d'avion de chasse d'où l'idée d'utiliser l'ouïe pour transmettre au pilote des informations concernant les menaces.

Qu'est ce que le son 3D?

Il s'agit de faire entendre, à l'aide d'un casque stereophonique, des sons ou des paroles qui soient perçus par l'auditeur comme issus d'un point particulier de l'espace. Le son 3D fait partie du concept de réalité virtuelle.

Une remarque peut être faite. L'écoute au casque stéréophonique n'est pas nouvelle. Toutefois, jusqu'à ce jour, l'écoute souffre de défauts importants :

- la localisation est imprécise.
- Le son est perçu à l'intérieur de la tête, il n'est pas spatialisé.
- L'image sonore se déplace avec la tête du sujet, rendant impossible le repérage d'une source par rapport à un référentiel fixe,

Ce sont là des défauts auxquels il faut impérativement remédier pour qu'un système son 3D existe.

Afin d'apprécier les possibilités du son 3D, nous avons décidé de mettre au point un outil permettant un son 3D de la meilleure qualité possible : simulation individualisée, qualité Haute-Fidelité pour les convertisseurs, le filtrage en temps réel et les écouteurs.

Dans notre exposé, nous présenterons d'abord les applications possibles du son 3D. Nous décrirons le principe d'un système 3D et rappellerons les paramètres physiques qui permettent la localisation des sons par l'homme. Nous exposerons alors nos travaux et nous les discuterons.

DOMAINES D'APPLICATION DU SON 3D (ref 3)

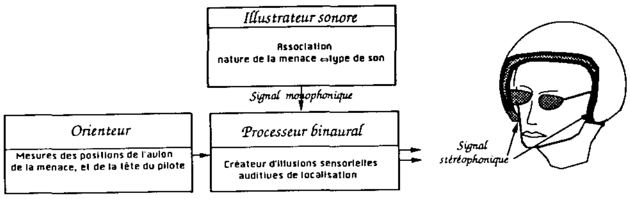
Le son 3D a deux grands domaines d'applications qui sont liés à la nature des informations à transmettre.

Le premier domaine est celui de la transmission des messages contenant essentiellement une information de position dans l'espace. Dans le cas d'un pilote d'avion de chasse, il peut s'agir de la position d'un missile qui le menace. La menace peut être également due à la trop grande proximité d'un avion ami. Le son 3D devient alors un système anticollision. Dans ce type d'application, il s'agit donc de transmettre le plus rapidement possible l' information de position.

Le système auditif est la voie idéale pour transmettre ce genre d'information, puisque finalement on ne fait qu'exploiter un réflexe de défense de l'homme. A l'écoute d'un son inquiétant, ou inhabituel, l'auditeur tourne sa tête dans la direction du son, afin de pouvoir affronter de face une éventuelle menace. Pour le pilote cela peut être un moyen rapide de lui indiquer où regarder.

Le deuxième grand domaine d'application concerne les cas où le message à transmettre est placé dans l'espace, afin de le rendre plus intelligible. Le son 3D permettra probablement d'améliorer les possibilités de surveillance simultanée de plusieurs communications audio en attribuant à chaque communication radio et à chaque avertissement audio une localisation dans une direction différente.

RÉALISATION D'UN SYSTÈME 3D (figure 1)



Principe du son 3D

Figure 1

La simulation 3D est décomposable en 3 fonctions :

L'illustrateur sonore.

A chaque type d'information à transmettre, il associe un son particulier. Il peur s'agir d'un son synthétisé. Il peut s'agir également d'une voix humaine.

L'Orienteur.

A partir des positions de l'avion, de la menace et de la tête du pilote, il calcule les coordonnées du point de l'espace où doit être localisé le signal sonore.

Le Processeur binaural.

Presented at an AGARD Meeting on 'Virtual Interfaces: Research and Applications', October 1993.

A partir d'un signal monophonique, fourni par l'illustrateur sonore, il fabrique deux signaux, un pour chaque oreille, tels qu'à l'écoute, l'auditeur localise le son dans l'espace " réel" à trois dimensions, à l'emplacement fourni par l'orienteur. Ce processeur est le coeur du système.

MÉCANISMES UTILISÉS PAR L'HOMME POUR DÉTERMINER LA DIRECTION D'UN SON (ref 2)

En premier lieu, il y a le retard interaural:

-Le signal issu d'une source sonore parvient à des instants différents aux deux oreilles.

- Les rotations éventuelles de la tête, même celles de petite amplitude et qui sont plus ou moins conscientes améliorent sensiblement les possibilités de repérage par une technique proche de la goniométrie.

En deuxième lieu, il y a le spectre des signaux.

- Le contenu fréquentiel des signaux parvenant aux deux oreilles est en général différent car la tête et le torse font obstacle aux ondes sonores. Cet effet d'obstacle dépend bien entendu de la fréquence.

-Les pavillons des oreilles réalisent également un filtrage (ref 1).

On appelle fonction de transfert de tête, le rapport des transformées de fourier des pressions acoustiques à l'emplacement de la tête du sujet , avec et sans le sujet. La figure 2 présente un exemple de ces caractéristiques relevées dans le plan horizontal, sur une tête artificielle. Ce filtrage, notons le, varie d'un individu à l'autre, notamment du fait de la diversité de forme des pavillons des oreilles. L'effet est surtout sensible sur la localisation en site (ref6). C'est pourquoi nous avons décidé de mesurer individuellement, en 3D, les caractéristiques du filtrage acoustique réalisé par le thorax, la tête et les pavillons d'un sujet.

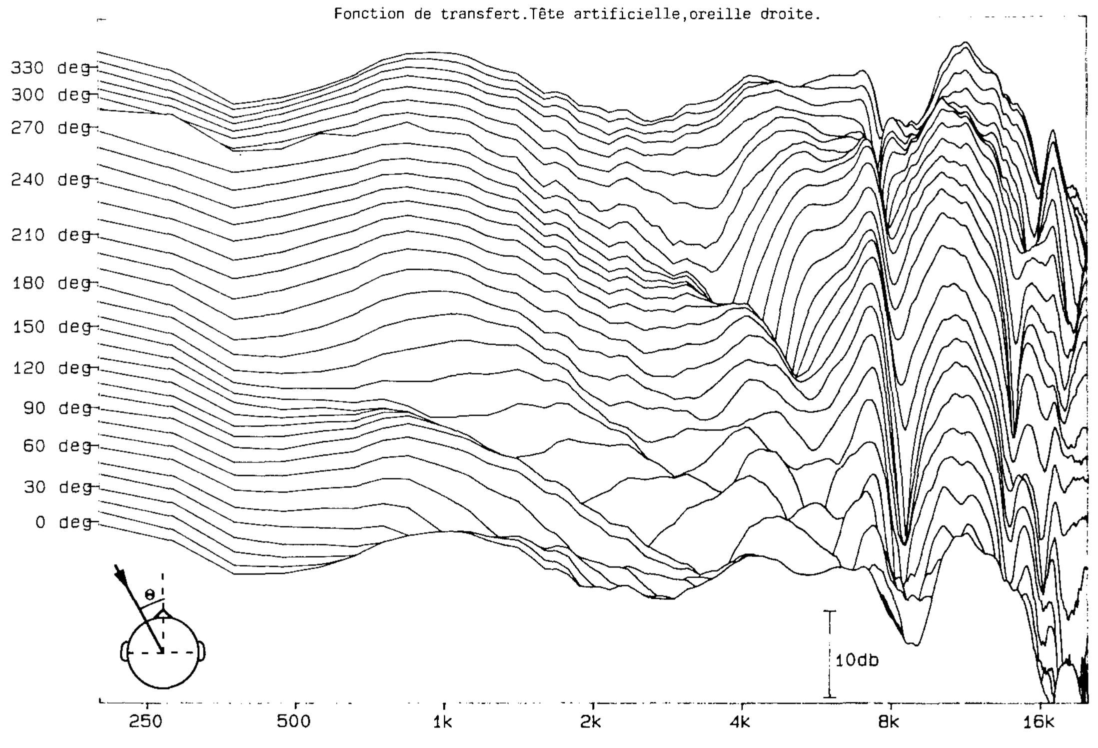


Figure 2

TRAVAUX RÉALISÉS

- La première partie de notre travail a été consacrée à la mesure des caractéristiques de 4 sujets.

- La deuxième partie a été consacrée aux test d'écoute. Il est demandé au sujet de localiser des sons synthétisés soit à partir de ses propres caractéristiques, soit à partir de caractéristiques d'un autre sujet. Autrement dit, les sujets écoutent soit avec leurs pavillons, soit avec ceux d'un autre !

Technique de mesure.

Cette technique est inspirée de celles utilisées par Pösselt en 87 (ref 4) et Wightman en 89 (ref 7).

Nous avons pris une empreinte des conduits auditifs des sujets. Nous avons ensuite réalisé sur mesure des bouchons contenant les microphones. Les bouchons sont insérés dans les conduits auditifs qui sont donc totalement obstrués.

La photo 1 présente le dispositif de mesure. Un Haut-parleur se déplace sur un rail, de façon à explorer n'importe quel point d'une sphère centrée sur la tête d'un sujet assis au centre de la pièce. Le processus est entièrement automatisé

Le signal d'excitation envoyé dans le haut-parleur est une séquence binaire pseudoaléatoire. Le champ acoustique émis par le haut parleur est cartographié une première fois, en l'absence du sujet. Le champ est à nouveau mesuré en présence du sujet, à l'aide des micros insérés dans les conduits auditifs. Le rapport des deux pressions acoustiques détermine la fonction de transfert du sujet. Cette technique élimine les caractéristiques du haut parleur et des

microphones. Ceci est rendu possible par une répétitivité très précise du positionnement du haut parleur puisque l'erreur est largement inférieure au degré.

La position absolue du sujet, en gisement, est fixée par une méthode acoustique en annulant le déphasage interaural lorsque le haut parleur est placé au gisement 0. Quant au site, c'est le sujet lui même qui détermine sa position horizontale. Les positions absolues étant repérées, le sujet contrôle lui même la fixité de sa posture grâce à un système de deux cameras video et à l'aide de marqueurs video de face et de profil.

Les mesures sont faites en trois séances consécutives, durant chacune environ 1 demi heure. Au total on a une base de données de 450 paires de fonctions de transfert par sujet.

Les réponses impulsionnelles correspondantes sont ensuite chargées dans un PC équipé d'un processeur Convolvotron (ref 5). Il devient alors possible de réaliser le filtrage en temps réel d'un signal audio. On fabrique ainsi à partir d'un signal monophonique un signal spatialisé à l'emplacement voulu par l'opérateur. C'est ce signal stéréo que l'on envoie dans le casque d'un sujet.

photo1

expérimentation psychoacoustique.
Nous avons décidé dans ces expériences préliminaires d'apprécier la précision qu'il est possible d'obtenir lorsqu'il est demandé à un sujet de faire face à une source sonore simulée à l'aide du casque.

Le signal sonore à localiser, une voix, est émis pendant tout le temps de la recherche. Le sujet prend tout son temps pour répondre, car c'est la précision que l'on veut évaluer (en fait, le temps de recherche est limité à 30 secondes).

Plus précisément, le sujet doit viser la source à l'aide d'une croix collimatée solidaire du casque. Ainsi, la direction du regard est fixe par rapport au casque. Sur celui ci est fixé un détecteur de position électromagnétique Bird. Les indications du bird asservissent les signaux audio afin que la source sonore simulée garde une position constante dans le repère lié au local. Le bird a été soigneusement étalonné : les différents repères ont été harmonisés à l'aide de visées optiques par théodolites et d'un deuxième réticule collimaté à l'infini et placé dans le local.

Le sujet est assis au même endroit que dans la première partie, dans une semi obscurité, afin de minimiser le rôle des repères visuels..

Les tests sont limités à un ensemble de 12 points simulés, dont les gisements sont espacés de 30 degrés et dont les sites sont : 0°,+36° et -36°. Chaque emplacement est présenté 3 fois au cours du test, et de façon aléatoire. On a ainsi 36 réponses par test. Le test dure environ 20 minutes.

Les rotations de la tête sont relevées tuutes les 40ms et stockées. Cela permet l'analyse éventuelle des stratégies de recherche des sujets.

Notons que tous les résultats que nous présentons ici ont été obtenus lorsque les sujets écoutaient avec leurs propres pavillons .

Résultats de localisation en gisement (figure 3).
Chaque graphique correspond à un sujet. On a en abscisse le gisement réel, et en ordonnée les réponses du sujet. Les droites correspondent aux réponses idéales. Les résultats du haut sont bien regroupés autour de la droite idéale, l'erreur d'appréciation commise par les sujets est réduite. Ce sont les deux meilleurs sujets. En bas, c'est moins bien. Il y a des imprécisions, voire des erreurs.

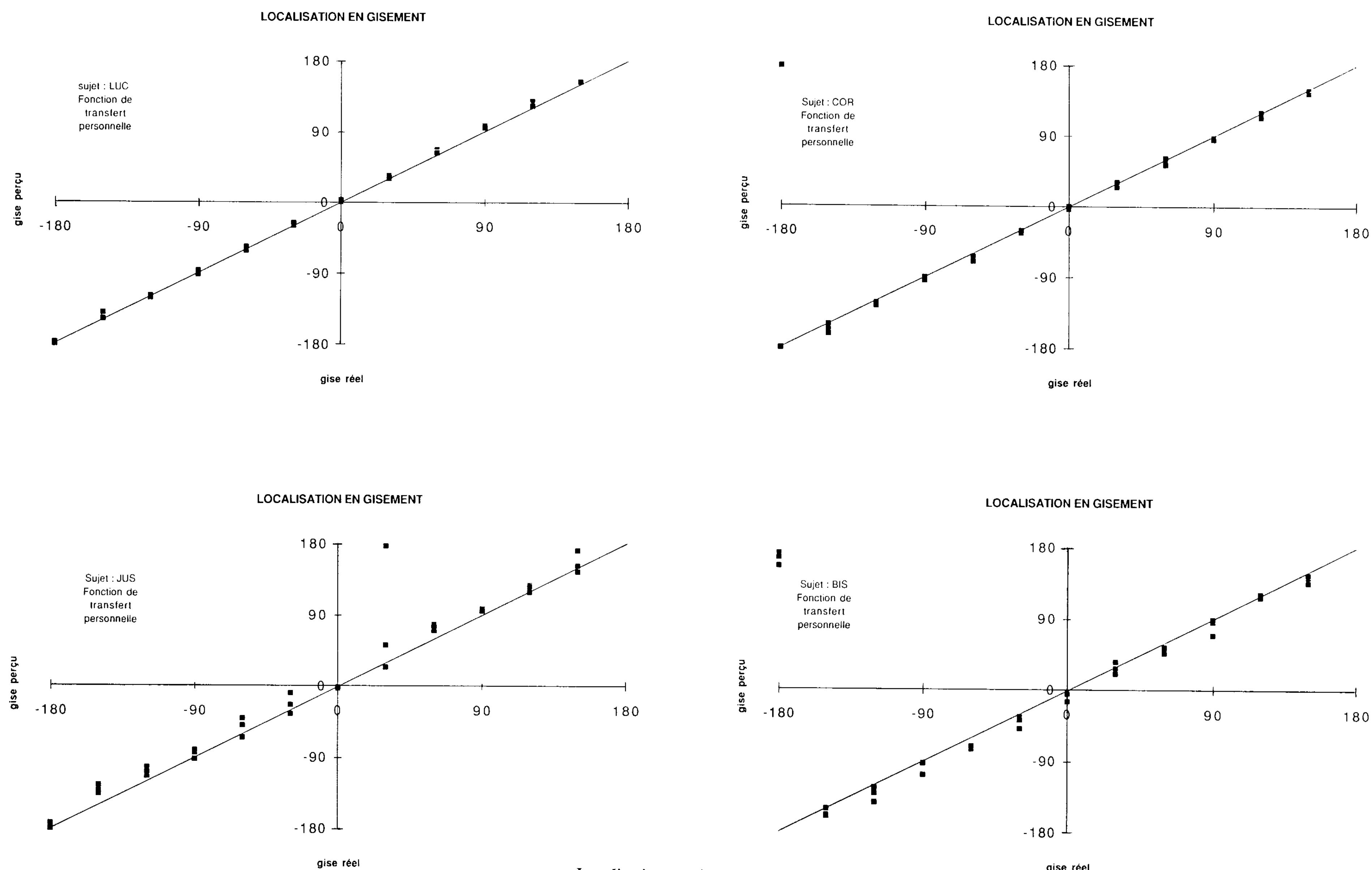
Meilleures performances en gisement (en degré)

Sujet	LUC	COR	JUS	BIS
Erreur moyenne	3	3	13	6
Erreur RMS	4	4	27	7
Ecart type	2	3	24	4
Erreur max	8	11	148	17

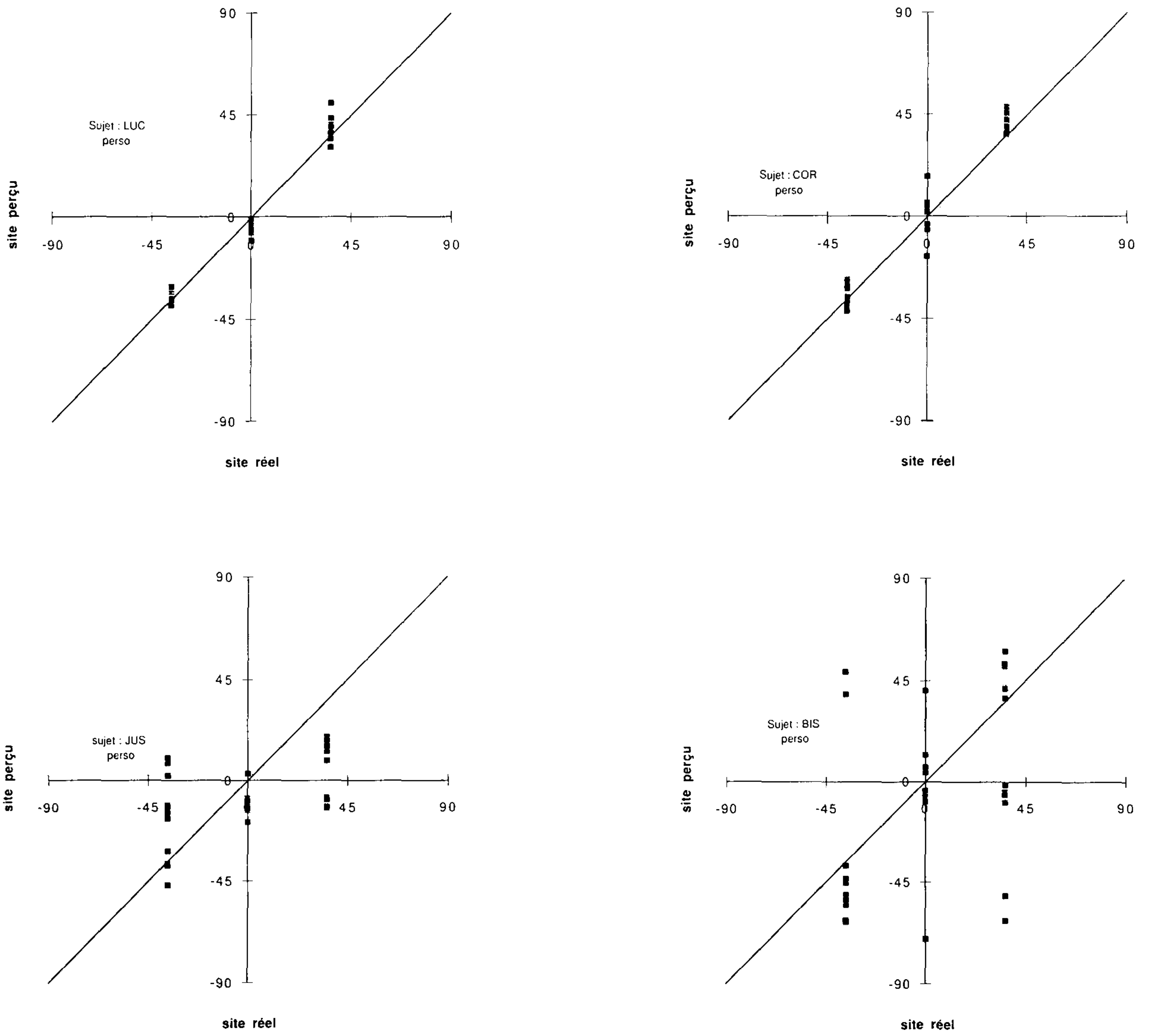
Tableau 1

Les valeurs statistiques sont présentées sur le tableau 1. L'erreur rms est de 4 degrés, pour les deux meilleurs, l'écart type est de 2 °ou 3 °. Les plus mauvaises performances correspondent à une erreur rms de 27° avec un écart type de 24°.

Résultats de localisation en site (figure 4).
Là encore, on retrouve les résultats de nos 4 sujets, disposés de la même façon. Les droites correspondent aux réponses idéales. Il n'y a que 3 séries de valeur puisque l'on n'a testé que 3 sites 0,+36° et - 36°. En haut, les résultats sont groupés et centrés autour des bonnes valeurs. Chez JUS, les points ne



Localisation en gisement. Figure 3



Localisation en site. Figure 4

sont pas trop dispersés, mais les valeurs centrales surtout à +36° sont décalées. Les sons venant du haut sont alors perçus moins haut qu'ils ne le sont. Notre dernier sujet, qui était le moins entrainé semble répondre de façon aléatoire.

Meilleures performances en site (en degré)

Sujet	LUC	COR	JUS	BIS
Erreur moyenne	4	6	20	25
Erreur RMS	5	7	25	36
Ecart type	3	5	14	27
Erreur max	14	18	48	98

Tableau 2

Les valeurs statistiques sont présentées sur le tableau 2 L'erreur rms est 5° et 7 ° avec des écarts type de 3 et 5° pour les 2 meilleurs.

Meilleures performances en ligne de visée (en degré)

Sujet	LUC	COR	JUS	BIS
Erreur moyenne	5	6	26	26
Erreur RMS	6	8	33	37
Ecart type	3	5	21	26
Erreur max	15	18	118	98

Tableau 3

L'erreur sur la ligne de visée (tableau 3) est certainement le meilleur critère de localisation. Cette erreur correspond à l'angle formé par la direction visée par le sujet et la vraie direction de la source. Pour 2 sujets, l'erreur rms est de 6 et 8° avec un écart type de 3 et 5°. Pour les 2 autres sujets, les résultats sont moins bons : 26° avec un écart type d'une vingtaine de degrés.

DISCUSSION

Les résultats de nos deux meilleurs sujets, tels qu'ils sont présentés semblent être du même ordre. Et pourtant, ils correspondent à des stratégies de recherche du site tout à fait différentes. En effet deux stratégies se dégagent:

- une stratégie "visuelle". Pendant la recherche en site, le sujet conserve un gisement proche de celui de la source. Disons qu'il place la source dans son champ visuel.
- une stratégie "auditive". De façon délibérée, le sujet tend sa meilleure oreille vers la source. C'est la stratégie des aveugles. Une fois le site déterminé, le sujet vient se placer face à la source, ainsi que cela lui est demandé.

Le premier sujet, LUC, a une stratégie visuelle. Il a atteint ses meilleures performances dès la 2ème séance et ses mouvements de tête sont assez limités. COR par contre n'a atteint les performances indiquées qu'après avoir adopté la stratégie auditive et en faisant beaucoup d'efforts.

CONCLUSION

Mis à part le fait qu'il y a encore beaucoup de travail à fournir et de nombreux tests psychoacoustiques à réaliser, il apparait que si la localisation en gisement est facile, celle en site est nettement plus difficile. Certains sujets étant nettement plus aptes que d'autres, des problèmes de sélection vont apparaitre, cette sélection ne pouvant se faire à partir des tests audiométriques classiques. Quant à ces tests classiques, ils devront élargir la bande de fréquence analysée car une bonne localisation en site nécessite une audition intègre dans les hautes fréquences, au delà de la bande conversationnelle habituellement testée. Enfin, sachant que la sensibilité en haute fréquence est souvent la première atteinte lors de l'exposition au bruit, voilà encore une raison pour renforcer la protection du personnel.

REMERCIEMENTS

Les travaux présentés ont été réalisés dans le cadre d'une étude financée par la DCAE / STTE / AVI (Direction des Constructions Aéronautiques, Service Technique des Télécommunications et des Equipements aéronautiques).

BIBLIOGRAPHIE

1 Batteau D.W. . The role of the pinna in human localisation. Proceedings of the royal society of London,Series B,168, pp 158-180 (1967)

2 Blauert J. .Spatial hearing. MIT Press, Cambridge, MA (1983)

3 Doll T.J.:Development of simulated directionnal audio for cockpit applications. rapport AAMRL TR 86-014 (1986).

4 Posselt C. Schroter J. Opitz M.,Divenyi P.L. , Blauert J. . Generation of binaural signals for research and home entertainment. 12ème congrès d'acoustique. B1-6 Toronto (1986).

5 Wenzel E.M., Wightman F.L., Foster S.H. A virtual display system for conveying three-dimensionnal acoustic information. Proceedings of the human factors society 32nd annual meeting(pp86-90).Santa Monica,CA:Human Factors society.

6 Wenzel E. ,Wightman F. , Kistler D. , Foster S.:Acoustic origins of individual differences in soun Suppl 1 Vol 84, Fall1988.

7 Wightman F.L. Kistler D.J. . Headphone simulation of free field listening. 1:Stimulus synthesis.J.Acoust.Soc.Am.85(2)pp858-867 February 1989

THE USE OF A TACTILE INTERFACE TO CONVEY POSITION AND MOTION PERCEPTIONS

A. H. Rupert [1,2]
F. E. Guedry [1,3]
M. F. Reschke [2]

[1] Naval Aerospace Medical Research Laboratory
Acceleration Division
51 Hovey Road
Pensacola, Florida 32508-1046

[2] NASA Johnson Space Center (SDS)
Space Biomedical Research Institute
Mail Code SD-5
Houston, Texas 77058-1113

[3] University of West Florida
Department of Psychology
Building 13, Room 201
11000 University Parkway
Pensacola, Florida 32514
USA

SUMMARY

Under normal terrestrial conditions, perception of position and motion is determined by central nervous system integration of concordant and redundant information from multiple sensory channels (somatosensory, vestibular, visual), which collectively yield veridical perceptions. In the acceleration environment experienced by pilots, the somatosensory and vestibular sensors frequently present false information concerning the direction of gravity. When presented with conflicting sensory information, it is **normal** for pilots to experience episodes of disorientation.

We have developed a tactile interface that obtains veridical roll and pitch information from a gyro-stabilized attitude indicator and maps this information in a one-to-one correspondence onto the torso of the body using a matrix of vibrotactors. This enables the pilot to continuously maintain an awareness of aircraft attitude without reference to visual cues, utilizing a sensory channel that normally operates at the subconscious level. Although initially developed to improve pilot spatial awareness, this device has obvious applications to 1) simulation and training, 2) nonvisual tracking of targets, which can reduce the need for pilots to make head movements in the high-G environment of aerial combat, and 3) orientation in environments with minimal somatosensory cues (e.g., underwater) or gravitational cues (e.g., space).

INTRODUCTION

In our day-to-day terrestrial activities, position and motion perception is continuously maintained by accurate information from three independent, overlapping, and concordant sensory systems: the visual, the vestibular (or inner ear), and the somatosensory systems (skin, joint and muscle sensors). These complementary and reliable sources of information are integrated in the central

Presented at an AGARD Meeting on 'Virtual Interfaces: Research and Applications', October 1993.

nervous system to help the organism formulate an appropriate motor response. The relative contribution of the various senses involved in the perception of one's orientation can be significantly altered by exposure to unusual sensory environments, resulting in perceptions that are no longer veridical.

For example, somatosensory pressure cues are markedly reduced underwater, while in the military aviation environment, the almost continuous changes in acceleration and direction of aircraft motion expose aircrew to a resultant gravitoinertial force that is constantly changing in magnitude and direction. Under such circumstances, somatosensory and vestibular information concerning the direction of "down" may be incorrect, and increased reliance must be placed on visual information. Unfortunately, varying gravitoinertial force fields can also produce visual illusions of motion and position. Thus, in unusual sensory environments, the central nervous system has the added responsibility of determining which sensory information is valid.

Understandably, the typical spatial disorientation mishap occurs when the visual orientation system is compromised (e.g., temporary distraction, increased workload, transitions between visual and meteorological conditions, or reduced visibility). The central nervous system must then compute orientation with the remaining vestibular and somatosensory information that is at its disposal, however, this information is frequently incorrect. It is no wonder that spatial orientation is markedly impaired in the underwater and aerospace environments. Indeed, it is a physiologically normal response to experience spatial disorientation in such circumstances. Virtual reality displays offer the opportunity to "correct" the position and motion illusions that occur in unusual acceleration and proprioceptive environments.

Current simulators and virtual reality devices use visual displays to adequately convey the perception of static position and attitude. However, a veridical awareness of dynamic motions (e.g., changes in attitude, velocity, and acceleration) is inadequately maintained by visual information alone and requires the addition of proprioceptive (somatosensory and vestibular) cues. For example, motion-based simulators attempt to simulate maintained acceleration either by utilizing 1) transient linear acceleration with "washout" of motion cues accompanied by visual representation of acceleration, or 2) change of pitch (tilt) to convey prolonged linear acceleration. Both methods possess inherent deficiencies. The former method is restricted by the limited linear travel available in current simulators, and in the latter method, linear motion perception is "contaminated" with the unavoidable canal stimulus produced in affecting a change in pitch. The current models of perception are capable of predicting responses for simple conditions of static vision and constant acceleration, however, the experiments required to extend the model to include the dynamic conditions experienced in aviation have not yet been carried out.

Cutaneous sensory information is not currently used to provide position or motion information to pilots. We propose that spatial orientation can be continuously maintained by providing information from the aircraft attitude sensor to the pilot through the nonutilized sensory channel of touch (Rupert, Mateczun, and Guedry, 1990).

One approach is to use a torso harness fitted with multiple electromechanical tactors that can continuously update the pilot's awareness of position. This is analogous to the way our brain obtains orientation in the terrestrial environment. Thus, the pilot should be able to maintain orientation information in the absence of a visual horizon or during inevitable gaze shifts from the aircraft instrument panel. This device should free the pilot to devote more time to weapons delivery systems and other tasks requiring visual attention.

The rationale for utilizing touch to convey position and motion perception and to overcome vestibular, visual, and auditory

illusions produced by unusual acceleration environments is based largely on knowledge about the ontology of sensory development (Fig. 1). In most vertebrates, the proprioceptive tactile system is the first sensory system to develop, followed by the vestibular system, then the auditory system, and finally the visual system (Gottlieb, 1971). In fact, the proprioceptive systems of somatosensory and vestibular function develop a rich interaction *in utero*. This follows logically since the somatosensory system needs information very early in development concerning the direction of the gravity vector in order to properly control the antigravity and gravity muscles. It is only much later in development that the auditory and visual systems are integrated into this already well-functioning proprioceptive system. The primacy of touch and somatosensation in the development of orienting behavior has been demonstrated in several neurophysiological and anatomical studies (Meredith and Stein, 1986a,b). We propose that by providing veridical somatosensory orientation information the illusory effects present in unusual acceleration and proprioceptive environments can be reduced to improve human performance in many facets of the military theater.

PROCEDURES AND INITIAL OBSERVATIONS

We constructed a series of prototype devices to determine whether a pilot could maintain normal orientation and control over an aircraft using tactile cues. Multiple tactors were placed on a torso suit to represent all directions of roll and pitch (Fig. 2). An aircraft attitude sensor (Fig. 3) provided roll and pitch information to an IBM portable computer that selected (via circuit designed in-house) the appropriate pattern of stimulation for a given combination of pitch and roll. The circuit takes data from an IBM (or compatible) PC parallel printer port and expands it to drive a matrix of stimulators (Cushman, 1993). Although the current matrix is an 8 x 24 (146 element) array, the user can define a maximum of 16 x 16 elements.

The tactors in the prototype display are miniature electromechanical speakers 1/8 of an inch in thickness and 1 inch in diameter. The stimulus waveform consists of 10 pulses of a 150 Hz rectangular pulse train operating at a 10% duty cycle, followed by a break of approximately 450 ms. A stretch lycra suit maintains an appropriate interface pressure between the tactor and the skin. The software programs to drive the tactors evolved continuously in response to feedback from each user and have been tailored to meet the requirements of each community (e.g., attitude awareness for aircraft vice vehicle velocity information for diving submersibles).

One program that we developed and tested in an aircraft conveys the direction of "down," or the gravity vector. To experience in the laboratory the sensation of roll and/or pitch as presented to the pilot, the gyro-attitude sensor was replaced with a joystick, which permitted subjects to experience the same tactile sensation on the ground and visually observe the equivalent aircraft orientation changes represented on the computer by standard aircraft attitude indicator symbology.

In this configuration, most subjects could learn within 30 min how to ascertain within 5 deg the pitch and roll information presented on their torso display. The pitch and roll limits of the current display are ±15 deg and ±45 deg respectively. Alternatively, subjects using the device in a closed-loop configuration could position by tactile cues alone the simulated attitude of the aircraft to within 5 deg of accuracy in pitch and roll. Similar accuracies were attained in actual flights in aircraft with no reference to instruments or outside visual cues.

When used to convey constant velocity, the column in the direction of the desired perception was stimulated first and followed by sequential activation of the three paired columns to each side of the first column stimulated. The perception was similar to the feeling of directed flow of fluid over the torso with the direction defined by the first column stimulated and the velocity

determined by the stimulus interval between activation of column pairs.

DISCUSSION

Preliminary results both in aircraft and in the laboratory indicate that orientation awareness can be maintained nonvisually in environments known to produce spatial disorientation. This device has many applications in addition to aircraft control.

When used in aircraft simulators (either with or without a motion base) to indicate attitude and changes in attitude or velocity, it will provide the pilot or trainee with additional cues that can be used for aircraft control, when transferring to the aircraft.

Studies by the U.S. Army (Simmons, Lees, and Kimbal, 1978a,b) indicate that pilots in instrument flight conditions spend more than 50 % of their visual scan time attending to two instruments, the ADI/attitude indicator and the directional gyro. By presenting this information nonvisually, pilots will be free to attend to other tasks and instruments that do require vision attention. Thus, not only will the introduction of spatial orientation information offered by tactile cues reduce spatial disorientation mishaps, it will also improve the mission effectiveness of the operator.

A person who is tapped by another on the shoulder or torso reflexively turns his attention to the area stimulated. The torso suit can take advantage of this basic reflex to direct attention to any target that is on sonar or radar or is being electronically tracked. Currently, for pilots, naval flight officers, or radar operators to acquire targets, they must devote their attention to the radar screen, cognitively ascertain the direction to which to direct their gaze, and then carry out the motor act of acquiring the target. In high workload environments with multiple targets, it is possible to represent one or possibly more targets tactually and aid pilots/operators in the rapid identification of friend or foe. An increasingly prevalent training technique in the military is interactive simulation of war theaters with multiple pilots engaged in the same dogfight, but with each in their own simulator. Tactile cues will improve the situational awareness of all participants. This device has obvious applications to personnel in command and control centers who need to maintain an awareness of geographic location of incoming information from a wide variety of platforms (ship, aircraft, tanks, infantry, satellite systems, etc.) from multiple geographic sites. Civilian applications include air traffic controllers and dispatchers.

Space applications to improve astronaut performance fall into several categories. In space, astronauts are deprived of any constant proprioceptive reference to indicate down. An appropriate model of our torso suit could be interfaced with an accurate inertial platform reference to give astronauts a continuous perception of orientation at a low level of awareness, in a manner analogous to the situation on Earth. In extra vehicular activities (EVA), this device could be used to present a constant point of reference, such as the floor of the space shuttle, the position of a satellite or telescope being repaired, or even the direction of the center of the earth. During EVA activities, the only sensory indication of velocity is currently provided by vision. Using the velocity presentation mode mentioned earlier, the astronaut can have indications of motion as good or even better that those available on Earth. This device can thus overcome the limitations of the vestibular system, which does not detect constant rotation or constant velocity, and instead of providing "virtual" reality, can go one step further to "hyper-reality" or better than normal maintenance of orientation.

This device offers a countermeasure to the sensory-motor disorders associated with adaptation to the microgravity conditions in space and readaptation to 1 G on returning from earth orbit. Space motion sickness, or space adaptation syndrome, has been attributed either wholly or in part to a rearrangement of proprioceptive sensory information, especially the absence of continuous vestibular otolith stimulation and reduced somatosensory information. Some astronauts have indicated that it is discomforting on entering orbit to lose the

sense of orientation awareness. By providing astronauts with a somatosensory reference to "down," and an enhanced awareness of motion (velocity information), they should be able to maintain an accurate perception of position and motion during transition periods.

Finally, the problem of sensory-motor incoordination on return to 1 G may be reduced by providing continuous position and motion cues throughout the mission. For example, we may be able to attenuate the otolith tilt-translation reinterpretation effect (Parker, 1985) that develops in space and disturbs an astronaut's sensory-motor readaptation to 1 G upon reentry. The torso suit would provide accurate somatosensory information that would not confirm the troublesome vestibular signals that cause this effect. Thus, we would expect that if astronauts were trained to attend to this reliable source of information, the magnitude of this problem could be significantly reduced.

There are a variety of well-known tactile illusions that will serve to enhance the effectiveness of tactile torso and limb devices. When using multiple stimulators, as in the torso vest, it is possible to take advantage of basic psychophysical principles to effect changes in perceived magnitude, position, and motion of the stimulus. For example, the perceived magnitude of a pair of vibrotactile stimuli presented in close temporal succession is dependent on the relative frequencies of the two stimuli (Verrillo and Gescheider, 1975). The perceived position of two tactile pulses presented in rapid succession at different spatial locations will appear as a single moving source (Kirman, 1974). The latter principle was used in the torso suit to create the sensation of motion or directional flow over the thorax. Using these and other sensory illusions, it is actually possible to create compelling position and motion perceptions using fewer tactors than in our current prototype suit. Hans-Leukas Teuber (1956) demonstrated that the number of dimensions of a perception will exceed that of the physical stimuli. By varying only the frequency and intensity, his subjects experienced changes in pitch, density, volume, and loudness. Given the large number of available stimulus parameters (intensity, frequency, body position, interstimulus interval, multiple tactors, etc.), it will be possible to tactually present a wide variety of perceptions to convey position, motion, and target information in a way analogous to the observations of Teuber.

This tactile interface will contribute to basic research concerning haptic contributions to the interaction and integration of sensory information and vestibular brainstem reflexes, as well as the perceptual phenomena perceived at the cortical level. Inclusion of the haptic component will permit us to further refine and extend our model of sensory-motor interaction. The ultimate practical goal is to provide accurate predictive information to enhance the effectiveness of human factors engineers in the design of improved man-machine interfaces.

ACKNOWLEDGMENTS

This research was sponsored in part by the Naval Medical Research and Development Command under project number 63706N M0096.002-7056.

The views expressed in this article are those of the authors and do not reflect the official policy or position of the Department of Defense, the Department of the Navy, nor the U. S. Government.

Volunteer subjects were recruited, evaluated, and employed in accordance with the procedures specified in the Department of Defense Directive 3216.2 and Secretary of the Navy Instruction 3900.39 series. These instructions are based on voluntary informed consent and meet or exceed the provision of prevailing national and international guidelines.

REFERENCES

Cushman, W. B. 1993. A parallel printer port to matrix driver with high current DAC output. Behavior Research Methods, Instruments, and Computers. 25 (1): 48-52.

Gottlieb, G. 1971. Ontogenesis of sensory function in birds and mammals. In: Tobach, E., Avonsen, L. R., and Shaw, E. (Eds.) The Biopsychology of Development. New York: Academic Press.

Kirman, J. H. 1974. Tactile apparent movement: The effects of interstimulus-onset interval and stimulus duration. Perception and Psychophysics. 15: 1-6.

Meredith, M. A. and Stein, B. E. 1986a. Visual, auditory, and somatosensory convergence on cells in superior colliculus results in multisensory integration. Journal of Neurophysiology. 56: 640-662.

Meredith, M. A. and Stein, B. E. 1986b. Spatial factors determine the activity of multisensory neurons in cat superior colliculus. Brain Research. 365: 350-354.

Parker, D. E, Reschke, M. F., Arrott, A. P., Homick, J. L., and Lichtenberg, B. K. 1985. Otolith Tilt-Translation Reinterpretation Followin Prolonged Weightlessness: Implications for Preflight Training. Aviation, Space, and Environmental Medicine. 56: 601-6.

Rupert, A. H., Mateczun, A. J., and Guedry, F. E. 1990. Maintaining Spatial Orientation Awareness. AGARD CP-478; Symposium at Copenhagen, Denmark, 2-6 October, 1989.

Simmons, R. R., Lees, M.A., and Kimball, K.A. 1978a. Visual Performance/Workload of Helicopter Pilots During Instrument Flight. In: AGARD Operational Helicopter Aviation Medicine.

Simmons, R. R., Lees, M.A., and Kimball, K. A. 1978b. Aviator Visual Performance: A Comparative Study of a Helicopter Simulator and the UH-1 Helicopter. In: AGARD Operational Helicopter Aviation Medicine.

Teuber, H.-L., and Liebert, R. S. 1956. Auditory Vection. Am. Psychol. 11: 430.

Turner, D. C., and Bateson, P., (Eds.) 1988. The Domestic Cat: The Biology of Its Behaviour. Cambridge: Bath Press.

Verrillo, R. T., and Gescheider, G. A. 1975. Enhancement and summation in the perception of two successive vibrotactile stimuli. Perception and Psychophysics. 18: 128-136.

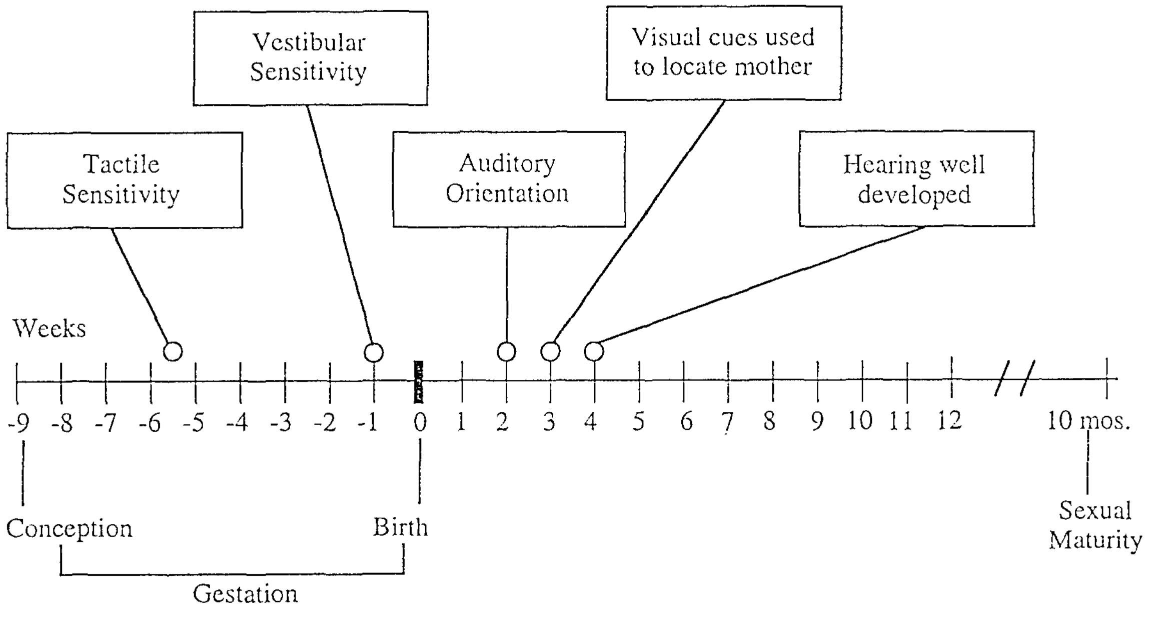


Figure 1 Timetable outlining sensory development of the domestic cat. (Turner and Bateson, 1988)

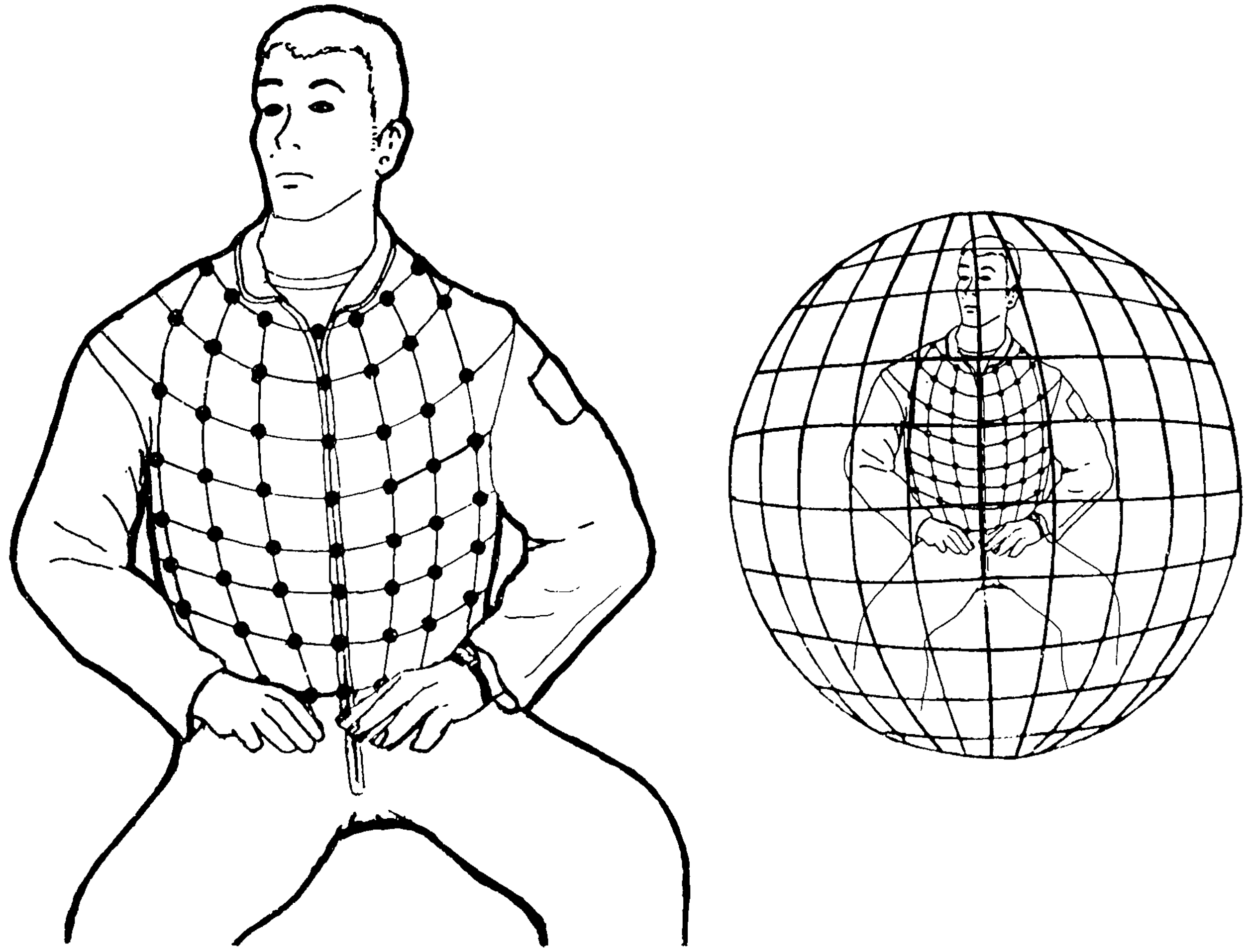


Fig 2 a) Tactor placement on torso vest with b) grid replacement of external environment superimposed.

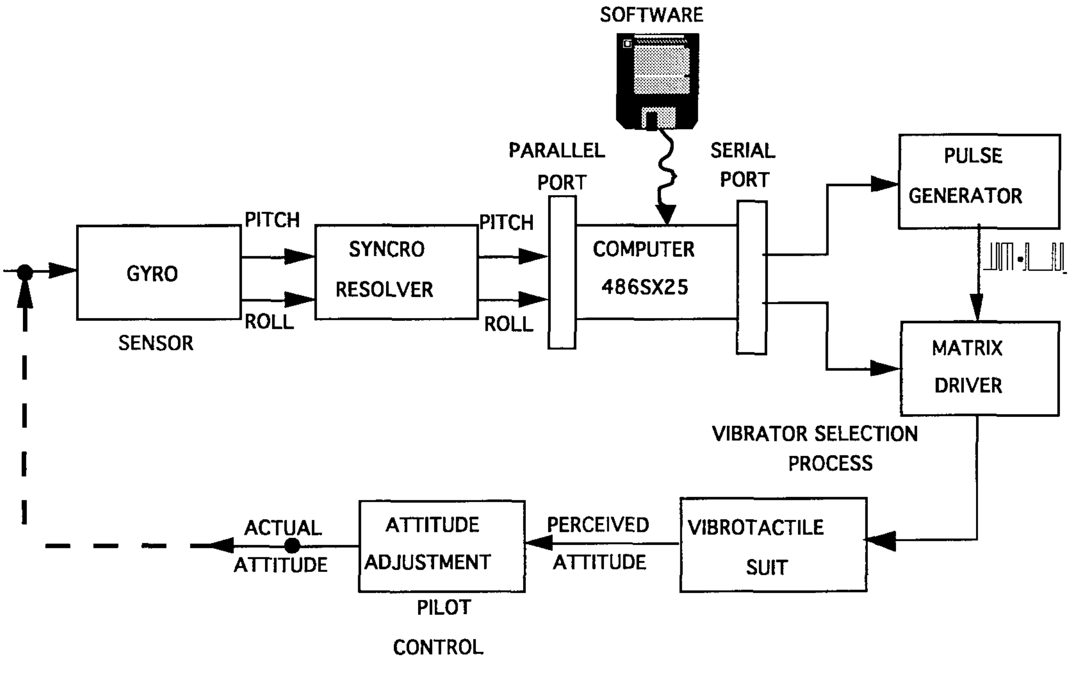


Figure 3 Schematic illustrating components of prototype tactile orientation system.

REPORT DOCUMENTATION PAGE

1. Recipient's Reference	2. Originator's Reference	3. Further Reference	4. Security Classification of Document
	AGARD-CP-541	ISBN 92-835-0746-0	UNCLASSIFIED/ UNLIMITED

5. Originator Advisory Group for Aerospace Research and Development
North Atlantic Treaty Organization
7 rue Ancelle, 92200 Neuilly sur Seine, France

6. Title VIRTUAL INTERFACES: RESEARCH AND APPLICATIONS

7. Presented at the Aerospace Medical Panel Symposium held in Lisbon, Portugal, 18th – 22nd October 1993.

8. Author(s)/Editor(s)	9. Date
Various	May 1994

10. Author(s)/Editor's Address	11. Pages
Various	192

12. Distribution Statement There are no restrictions on the distribution of this document. Information about the availability of this and other AGARD unclassified publications is given on the back cover.

13. Keywords/Descriptors

Human systems	Virtual technologies	Balance
Aerospace medicine	Simulation	Spatial orientation
Human factors engineering	Telerobotics	Situational awareness
Sensory processing	Vision	Cockpits
Display devices	Hearing	Pilots
Human interface	Haptics	Training

14. Abstract

Recent advances in technologies for information display and sensing of human movements, combined with computer based models of natural and artificial environments, have led to the introduction of so-called virtual interfaces. Virtual interfaces offer increased flexibility and naturalness, so are considered for use in several domains including aviation, training, design, simulation.

Papers presented at this symposium considered issues of research and application in virtual interfaces broadly defined. Issues of technology integration for system development were considered separately from issues of movement monitoring or sensory display. Issues of human performance measurement were presented in the context of both research and application.

AGARD Conference Proceedings 541
Advisory Group for Aerospace Research and Development, NATO
VIRTUAL INTERFACES:
RESEARCH AND APPLICATIONS
Published May 1994
192 pages

Recent advances in technologies for information display and sensing of human movements, combined with computer based models of natural and artificial environments, have led to the introduction of so-called virtual interfaces. Virtual interfaces offer increased flexibility and naturalness, so are considered for use in several domains including aviation, training, design, simulation.

P.T.O.

AGARD-CP-541

Human systems
Aerospace medicine
Human factors engineering
Sensory processing
Display devices
Human interface
Virtual technologies
Simulation
Telerobotics
Vision
Hearing
Haptics
Balance
Spatial orientation
Situational awareness
Cockpits
Pilots
Training

AGARD Conference Proceedings 541
Advisory Group for Aerospace Research and Development, NATO
VIRTUAL INTERFACES:
RESEARCH AND APPLICATIONS
Published May 1994
192 pages

Recent advances in technologies for information display and sensing of human movements, combined with computer based models of natural and artificial environments, have led to the introduction of so-called virtual interfaces. Virtual interfaces offer increased flexibility and naturalness, so are considered for use in several domains including aviation, training, design, simulation.

P.T.O.

AGARD-CP-541

Human systems
Aerospace medicine
Human factors engineering
Sensory processing
Display devices
Human interface
Virtual technologies
Simulation
Telerobotics
Vision
Hearing
Haptics
Balance
Spatial orientation
Situational awareness
Cockpits
Pilots
Training

AGARD Conference Proceedings 541
Advisory Group for Aerospace Research and Development, NATO
VIRTUAL INTERFACES:
RESEARCH AND APPLICATIONS
Published May 1994
192 pages

Recent advances in technologies for information display and sensing of human movements, combined with computer based models of natural and artificial environments, have led to the introduction of so-called virtual interfaces. Virtual interfaces offer increased flexibility and naturalness, so are considered for use in several domains including aviation, training, design, simulation.

P.T.O.

AGARD-CP-541

Human systems
Aerospace medicine
Human factors engineering
Sensory processing
Display devices
Human interface
Virtual technologies
Simulation
Telerobotics
Vision
Hearing
Haptics
Balance
Spatial orientation
Situational awareness
Cockpits
Pilots
Training

AGARD Conference Proceedings 541
Advisory Group for Aerospace Research and Development, NATO
VIRTUAL INTERFACES:
RESEARCH AND APPLICATIONS
Published May 1994
192 pages

Recent advances in technologies for information display and sensing of human movements, combined with computer based models of natural and artificial environments, have led to the introduction of so-called virtual interfaces. Virtual interfaces offer increased flexibility and naturalness, so are considered for use in several domains including aviation, training, design, simulation.

P.T.O.

AGARD-CP-541

Human systems
Aerospace medicine
Human factors engineering
Sensory processing
Display devices
Human interface
Virtual technologies
Simulation
Telerobotics
Vision
Hearing
Haptics
Balance
Spatial orientation
Situational awareness
Cockpits
Pilots
Training

Papers presented at this symposium considered issues of research and application in virtual interfaces broadly defined. Issues of technology integration for system development were considered separately from issues of movement monitoring or sensory display. Issues of human performance measurement were presented in the context of both research and application.

Papers presented at the Aerospace Medical Panel Symposium, held in Lisbon, Portugal, 18th – 22nd October 1993.

ISBN 92-835-0746-0

Papers presented at this symposium considered issues of research and application in virtual interfaces broadly defined. Issues of technology integration for system development were considered separately from issues of movement monitoring or sensory display. Issues of human performance measurement were presented in the context of both research and application.

Papers presented at the Aerospace Medical Panel Symposium, held in Lisbon, Portugal, 18th – 22nd October 1993.

ISBN 92-835-0746-0

Papers presented at this symposium considered issues of research and application in virtual interfaces broadly defined. Issues of technology integration for system development were considered separately from issues of movement monitoring or sensory display. Issues of human performance measurement were presented in the context of both research and application.

Papers presented at the Aerospace Medical Panel Symposium, held in Lisbon, Portugal, 18th – 22nd October 1993.

ISBN 92-835-0746-0

Papers presented at this symposium considered issues of research and application in virtual interfaces broadly defined. Issues of technology integration for system development were considered separately from issues of movement monitoring or sensory display. Issues of human performance measurement were presented in the context of both research and application.

Papers presented at the Aerospace Medical Panel Symposium, held in Lisbon, Portugal, 18th – 22nd October 1993.

ISBN 92-835-0746-0

AGARD

NATO - OTAN

7 RUE ANCELLE · 92200 NEUILLY-SUR-SEINE

FRANCE

Télécopie (1)47.38.57.99 · Télex 610 176

DIFFUSION DES PUBLICATIONS AGARD NON CLASSIFIEES

Aucun stock de publications n'a existé à AGARD. A partir de 1993, AGARD détiendra un stock limité des publications associées aux cycles de conférences et cours spéciaux ainsi que les AGARDographies et les rapports des groupes de travail, organisés et publiés à partir de 1993 inclus. Les demandes de renseignements doivent être adressées à AGARD par lettre ou par fax à l'adresse indiquée ci-dessus. *Veuillez ne pas téléphoner.* La diffusion initiale de toutes les publications de l'AGARD est effectuée auprès des pays membres de l'OTAN par l'intermédiaire des centres de distribution nationaux indiqués ci-dessous. Des exemplaires supplémentaires peuvent parfois être obtenus auprès de ces centres (à l'exception des Etats-Unis). Si vous souhaitez reçevoir toutes les publications de l'AGARD, ou simplement celles qui concernent certains Panels, vous pouvez demander à être inclu sur la liste d'envoi de l'un de ces centres. Les publications de l'AGARD sont en vente auprès des agences indiquées ci-dessous, sous forme de photocopie ou de microfiche.

CENTRES DE DIFFUSION NATIONAUX

ALLEMAGNE
Fachinformationszentrum,
Karlsruhe
D-7514 Eggenstein-Leopoldshafen 2

BELGIQUE
Coordonnateur AGARD-VSL
Etat-Major de la Force Aérienne
Quartier Reine Elisabeth
Rue d'Evere, 1140 Bruxelles

CANADA
Directeur du Service des Renseignements Scientifiques
Ministère de la Défense Nationale
Ottawa, Ontario K1A 0K2

DANEMARK
Danish Defence Research Establishment
Ryvangs Allé 1
P.O. Box 2715
DK-2100 Copenhagen Ø

ESPAGNE
INTA (AGARD Publications)
Pintor Rosales 34
28008 Madrid

ETATS-UNIS
NASA Headquarters
Attention: CF 37, Distribution Center
300 E Street, S.W.
Washington, D.C. 20546

FRANCE
O.N.E.R.A. (Direction)
29, Avenue de la Division Leclerc
92322 Châtillon Cedex

GRECE
Hellenic Air Force
Air War College
Scientific and Technical Library
Dekelia Air Force Base
Dekelia, Athens TGA 1010

ISLANDE
Director of Aviation
c/o Flugrad
Reykjavik

ITALIE
Aeronautica Militare
Ufficio del Delegato Nazionale all'AGARD
Aeroporto Pratica di Mare
00040 Pomezia (Roma)

LUXEMBOURG
Voir Belgique

NORVEGE
Norwegian Defence Research Establishment
Attn: Biblioteket
P.O. Box 25
N-2007 Kjeller

PAYS-BAS
Netherlands Delegation to AGARD
National Aerospace Laboratory NLR
P.O. Box 90502
1006 BM Amsterdam

PORTUGAL
Força Aérea Portuguesa
Centro de Documentação e Informação
Alfragide
2700 Amadora

ROYAUME UNI
Defence Research Information Centre
Kentigern House
65 Brown Street
Glasgow G2 8EX

TURQUIE
Millî Savunma Başkanlığı (MSB)
ARGE Daire Başkanlığı (ARGE)
Ankara

Le centre de distribution national des Etats-Unis ne détient PAS de stocks des publications de l'AGARD.
D'éventuelles demandes de photocopies doivent être formulées directement auprès du NASA Center for Aerospace Information (CASI) à l'adresse suivante:

AGENCES DE VENTE

NASA Center for
Aerospace Information (CASI)
800 Elkridge Landing Road
Linthicum Heights, MD 21090-2934
United States

ESA/Information Retrieval Service
European Space Agency
10, rue Mario Nikis
75015 Paris
France

The British Library
Document Supply Division
Boston Spa, Wetherby
West Yorkshire LS23 7BQ
Royaume Uni

Les demandes de microfiches ou de photocopies de documents AGARD (y compris les demandes faites auprès du CASI) doivent comporter la dénomination AGARD, ainsi que le numéro de série d'AGARD (par exemple AGARD-AG-315). Des informations analogues, telles que le titre et la date de publication sont souhaitables. Veuiller noter qu'il y a lieu de spécifier AGARD-R-nnn et AGARD-AR-nnn lors de la commande des rapports AGARD et des rapports consultatifs AGARD respectivement. Des références bibliographiques complètes ainsi que des résumés des publications AGARD figurent dans les journaux suivants:

Scientifique and Technical Aerospace Reports (STAR)
publié par la NASA Scientific and Technical
Information Program
NASA Headquarters (JTT)
Washington D.C. 20546
Etats-Unis

Government Reports Announcements and Index (GRA&I)
publié par le National Technical Information Service
Springfield
Virginia 22161
Etats-Unis
(accessible également en mode interactif dans la base de données bibliographiques en ligne du NTIS, et sur CD-ROM)

Imprimé par Specialised Printing Services Limited
40 Chigwell Lane, Loughton, Essex IG10 3TZ

AGARD

NATO OTAN

7 RUE ANCELLE · 92200 NEUILLY-SUR-SEINE

FRANCE

Telefax (1)47.38.57.99 · Telex 610 176

DISTRIBUTION OF UNCLASSIFIED AGARD PUBLICATIONS

AGARD holds limited quantities of the publications that accompanied Lecture Series and Special Courses held in 1993 or later, and of AGARDographs and Working Group reports published from 1993 onward. For details, write or send a telefax to the address given above. *Please do not telephone.*

AGARD does not hold stocks of publications that accompanied earlier Lecture Series or Courses or of any other publications. Initial distribution of all AGARD publications is made to NATO nations through the National Distribution Centres listed below. Further copies are sometimes available from these centres (except in the United States). If you have a need to receive all AGARD publications, or just those relating to one or more specific AGARD Panels, they may be willing to include you (or your organisation) on their distribution list. AGARD publications may be purchased from the Sales Agencies listed below, in photocopy or microfiche form.

NATIONAL DISTRIBUTION CENTRES

BELGIUM
Coordonnateur AGARD — VSL
Etat-Major de la Force Aérienne
Quartier Reine Elisabeth
Rue d'Evere, 1140 Bruxelles

CANADA
Director Scientific Information Services
Dept of National Defence
Ottawa, Ontario K1A 0K2

DENMARK
Danish Defence Research Establishment
Ryvangs Allé 1
P.O. Box 2715
DK-2100 Copenhagen Ø

FRANCE
O.N.E.R.A. (Direction)
29 Avenue de la Division Leclerc
92322 Châtillon Cedex

GERMANY
Fachinformationszentrum
Karlsruhe
D-7514 Eggenstein-Leopoldshafen 2

GREECE
Hellenic Air Force
Air War College
Scientific and Technical Library
Dekelia Air Force Base
Dekelia, Athens TGA 1010

ICELAND
Director of Aviation
c/o Flugrad
Reykjavik

ITALY
Aeronautica Militare
Ufficio del Delegato Nazionale all'AGARD
Aeroporto Pratica di Mare
00040 Pomezia (Roma)

LUXEMBOURG
See Belgium

NETHERLANDS
Netherlands Delegation to AGARD
National Aerospace Laboratory, NLR
P.O. Box 90502
1006 BM Amsterdam

NORWAY
Norwegian Defence Research Establishment
Attn: Biblioteket
P.O. Box 25
N-2007 Kjeller

PORTUGAL
Força Aérea Portuguesa
Centro de Documentação e Informação
Alfragide
2700 Amadora

SPAIN
INTA (AGARD Publications)
Pintor Rosales 34
28008 Madrid

TURKEY
Milli Savunma Başkanlığı (MSB)
ARGE Daire Başkanlığı (ARGE)
Ankara

UNITED KINGDOM
Defence Research Information Centre
Kentigern House
65 Brown Street
Glasgow G2 8EX

UNITED STATES
NASA Headquarters
Attention: CF 37, Distribution Center
300 E Street, S.W.
Washington, D.C. 20546

The United States National Distribution Centre does NOT hold stocks of AGARD publications.
Applications for copies should be made direct to the NASA Center for Aerospace Information (CASI) at the address below.

SALES AGENCIES

NASA Center for
Aerospace Information (CASI)
800 Elkridge Landing Road
Linthicum Heights, MD 21090-2934
United States

ESA/Information Retrieval Service
European Space Agency
10, rue Mario Nikis
75015 Paris
France

The British Library
Document Supply Centre
Boston Spa, Wetherby
West Yorkshire LS23 7BQ
United Kingdom

Requests for microfiches or photocopies of AGARD documents (including requests to CASI) should include the word 'AGARD' and the AGARD serial number (for example AGARD-AG-315). Collateral information such as title and publication date is desirable. Note that AGARD Reports and Advisory Reports should be specified as AGARD-R-nnn and AGARD-AR-nnn, respectively. Full bibliographical references and abstracts of AGARD publications are given in the following journals:

Scientific and Technical Aerospace Reports (STAR)
published by NASA Scientific and Technical
Information Program
NASA Headquarters (JTT)
Washington D.C. 20546
United States

Government Reports Announcements and Index (GRA&I)
published by the National Technical Information Service
Springfield
Virginia 22161
United States
(also available online in the NTIS Bibliographic
Database or on CD-ROM)

SPS

Printed by Specialised Printing Services Limited
40 Chigwell Lane, Loughton, Essex IG10 3TZ